ISBN Number: 978-81-956864-2-1

3rd ISTE STUDENT SECTION CONVENTION 2026

ON

Transformational Changes in Technical Education through Artificial Intelligence

BIHAR & JHARKHAND SECTION

ACCREDITED WITH GRADE B++ NAAC



On

18th - 19th November, 2026

Organised by:

GGSESTC

GURU GOBIND SINGH EDUCATIONAL SOCIETY'S TECHNICAL CAMPUS

GGSESTC, KANDRA (V),
CHAS, BOKARO,
JHARKHAND - 827013

Host :

ISTE

(Bihar & Jharkhand Section)

Name of the Book : **3rd ISTE STUDENT SECTION CONVENTION 2026** on "Transformational Changes in Technical Education through Artificial Intelligence"
Name of Publisher : **Guru Gobind Singh Educational Society's Technical Campus**
Name of Editor : **Dr. Priyadarshi Jaruhar**

GGSESTC

GURU GOBIND SINGH EDUCATIONAL SOCIETY'S TECHNICAL CAMPUS

COLLEGE OF ENGINEERING AND MANAGEMENT

Approved by AICTE, Ministry of Education, Govt. of India, New Delhi, Affiliated to Jharkhand University of Technology, Ranchi

A Fifth Generation **Engineering College** with **NAAC** Accreditation

You are cordially invited to participate in

3rd ISTE STUDENT SECTION CONVENTION 2026

ON

Transformational Changes in Technical Education through Artificial Intelligence

BIHAR & JHARKHAND SECTION



On

18th - 19th November, 2026

Venue :

GGSESTC, Bokaro

Organised by:

GGSESTC
GURU GOBIND SINGH EDUCATIONAL SOCIETY'S TECHNICAL CAMPUS
GGSESTC, KANDRA (V),
CHAS, BOKARO, JHARKHAND - 827013

Host :

ISTE
(Bihar & Jharkhand Section)

RSVP :
Prof. (Dr.) Priyadarshi Jaruhar
(Principal Convener : 9822732264)

Phone No.: 06542 - 265 398 **Toll Free No.:** 18003455398

Title: Proceeding of ISTE Students' Convention 2026 (Bihar and Jharkhand Section)

Theme: Transformational changes in technical education through Artificial Intelligence

Author: Dr. Priyadarshi Jaruhar

Published by: Guru Gobind Singh Educational Society's Technical Campus

Publisher's Address: Kandra, Chas, Bokaro, Jharkhand-827013

Printed by:

Shivam Printers
Chas, Bokaro
Jharkhand-827013
India.

First Impression: 2026

Published on: 18th November, 2026

ISBN: 978-81-956864-2-1

Disclaimer

No part of this publication may be reproduce or transmitted in any Formby any means, electronic or mechanical, including photocopy, recording or any information storage and retrieval system , without permission in writing from the copyright owners

© GGSESTC, Kandra, Bokaro, 2026

Chief Patrons

- **Dr. Pratapsinh Kakasaheb Desai,** President, ISTE, New Delhi
- **Prof. (Dr.) D. K. Singh,** VC, Jharkhand University of Technology, Ranchi
- **Shri Tarsem Singh,** President, Guru Gobind Singh Educational Society, Bokaro

Patrons

- **Prof. (Dr.) S. M. Ali,** Executive Secretary, ISTE, New Delhi
- **Prof. (Dr.) Mithilesh Kumar Jha,** Principal, M.I.T Muzaffarpur
- **Prof. (Dr.) Pankaj Rai,** Director, BIT Sindri
- **Prof. (Dr.) Upendra Prasad,** Principal, Government Engineering College, Bokaro
- **Shri S.P. Singh,** Secretary, GGSES, Bokaro

Principal Convener

Dr. Priyadarshi Jaruhar, Director, GGSESTC, Bokaro

Convener

Prof. Alok Kumar, (Co-Ordinator ISTE Student Chapter) Asst. Prof., EEE, GGSESTC, Bokaro

Organizing secretaries

- **Prof. Goutam Kumar,** (Co-Ordinator ISTE Faculty Chapter) Asst. Prof., ECE, GGSESTC, Bokaro.
- **Dr. Manojit De**, IQAC coordinator, Assoc. Prof., BSH, GGSESTC, Bokaro.
- **Dr. Akash Arya**, Asst. Prof., HoD of EEE, GGSESTC, Bokaro

National advisory committee

- **Prof. Anil Kumar,** National Executive Council Member, ISTE, New Delhi.
- **Prof. Nishant Kumar,** Registrar (I/C), JUT, Ranchi.
- **Shri. Sneh Kumar,** Director (Curriculum Development), JUT, Ranchi.
- **Prof. (Dr.) A.W. Kharche,** Director, Padmashri Dr. V. B. Kolte College of Engg, Buldhana.
- **Prof. (Dr.) Anupam Kumar,** Professor & Dean, Textile Engg. MRSPTU, Bhatinda.
- **Prof. (Dr.) Maya Rajnarayan Ray,** Assoc. Prof., CE, BIT Sindri.
- **Dr. P. A. Khatwani, Professor,** Textile Engg., Sarvajanik College of Engg & Tech, Surat
- **Prof. (Dr.) Pradip Kumar Sadhu,** IIT (ISM), Dhanbad.
- **Prof.(Dr.) Sanjay,** Professor & Head of ME Dept. NIT Jamshedpur.
- **Dr. Ranveer Kumar,** Assoc. Prof. and HOD of the Dept. of Chemistry at the CIT, Ranchi.

Organizing Committee

- **Dr. Rajesh Kumar Tiwari,** Principal, RVSCET, Jamshedpur
- **Prof. (Dr.) L. Ranganathan,** Principal, CIT, Ranchi
- **Dr. R.P. Verma,** Assoc. Prof., CIVIL, GGSESTC, Bokaro
- **Dr. Yasir Karim,** Assoc. Prof., CIVIL, GGSESTC, Bokaro
- **Dr. A. P. Burnwal,** Assoc. Prof., BSH, GGSESTC, Bokaro
- **Prof. Salim Ahmad,** Asst. Prof., ECE, GGSESTC, Bokaro
- **Prof. Apurba Sinha,** Asst. Prof., CSE, GGSESTC, Bokaro
- **Prof. Md. Asif Faizi,** Asst. Prof., MBA, GGSESTC, Bokaro
- **Dr. Shafaque Rahmat,** Asst. Prof., MBA, GGSESTC, Bokaro
- **Prof. Shahnaaz Farhin,** Asst. Prof., BBA, GGSESTC, Bokaro
- **Dr. Shafaque Zaheer,** Asst. Prof., BSH/BBA, GGSESTC, Bokaro
- **Mrs. Anu Singh**, Asst. Librarian, GGSESTC, Bokaro
- **Dr. Deepak Kumar,** Asst. Prof., CSE, GGSESTC, Bokaro

- **Prof. Mahmood Alam,** Asst. Prof., ME, GGSESTC, Bokaro
- **Prof. Sumit Kumar Pandey,** Asst. Prof., ME, GGSESTC, Bokaro
- **Prof. Manoj Kumar,** Asst. Prof., ME, GGSESTC, Bokaro
- **Mr. Anil Kumar Singh,** System Administrator, GGSESTC, Bokaro

Technical Committee

- **Prof. (Dr.) Sanjay Kr. Singh,** Principal, Govt. Engg. College, Palamu
- **Shri Rajesh K. Singh,** CGM, BSL, Bokaro
- **Prof. (Dr.) A B Marathe,** Director, Shri Shivaji Education Society's Amravati's College of Engineering & Technology, Akola, (M. S.)
- **Dr. Deepak Kumar,** Asst. Prof., BSH, GGSESTC, Bokaro
- **Prof. Tej Bahadur Singh,** Asst. Prof., ME, GGSESTC, Bokaro
- **Adv. Pallavi Prasad,** Administrative Officer, GGSESTC, Bokaro
- **Dr. Sabita Kumari,** HR & Training Head, Twintech, ACI, Autodesk, Bhubaneswar
- **Mr. Vishwajeet Kumar Pathak,** System Administrator, GGSESTC, Bokaro
- All Principals of Engineering & Polytechnic institutions affiliated to JUT, Ranchi
- All HoDs, GGSESTC, Bokaro

Students Organizing Committee

- **Palak Kumari**, CSE, Member ISTE Student Chapter
- **Harshit Kriti Raj Mandal**, CSE, Member ISTE Student Chapter
- **Siddhi Kumari Choudhary**, CSE(DS) Member ISTE Student Chapter
- **Shambhavi Mishra**, CSE(DS), Member ISTE Student Chapter
- **Nishat Ahmad**, CSE, Member ISTE Student Chapter
- **Chandan Kumar**, BCA, Member ISTE Student Chapter
- **Kanchan Kumari**, CSE,(DS), Member ISTE Student Chapter
- **Samaya Fatima**, BBA, Member ISTE Student Chapter
- **Sahiba Akram**, BBA , Member ISTE Student Chapter
- **Sourav Kumar Paramanik**, CSE, Member ISTE Student Chapter
- **Neha Kumari Paul** ,EEE, Member ISTE Student Chapter
- **Siya Kumari**. EEE, Member ISTE Student Chapter
- **Sarthik Kumar**, ME, Member ISTE Student Chapter
- **Manas Mondal**, ME, Member ISTE Student Chapter
- **Shalu kumari**, MBA, Member ISTE Student Chapter
- **Manish Kumar Gupta**, CIVIL, Member, ISTE Student Chapter

Chief Editor

Dr. Priyadarshi Jaruhar, Director, GGSESTC, Bokaro

Associate Editor

Dr. Manojit De, IQAC coordinator, Assoc. Prof., BSH, GGSESTC, Bokaro

Editorial Board Members

- **Prof. (Dr.) Bikash Ghoshal**, Prof. & HoD of ME, GGSESTC, Bokaro
- **Prof. Abhilash Ghosh**, Asst. Prof., EEE, GGSESTC, Bokaro
- **Prof. Alok Kumar,** Asst. Prof., EEE, GGSESTC, Bokaro

Message from the Vice-Chancellor
Jharkhand University of Technology, Ranchi



I am delighted to extend my warm greetings and best wishes on the occasion of the 3rd ISTE Bihar & Jharkhand Student Convention 2026, centred on the theme *"Transformational Changes in Technical Education through Artificial Intelligence."*

We are witnessing a defining era in which Artificial Intelligence is reshaping the landscape of education, research, innovation and professional practice. Its integration into technical education calls for a progressive shift from conventional learning towards critical thinking, interdisciplinary knowledge, creativity and innovation.

While AI offers unprecedented opportunities, its true potential lies in how responsibly and effectively we harness it. Our educational institutions must prepare students not merely to adapt to technological advancements but to become innovators, problem-solvers and responsible leaders who can shape the future.

I believe that platforms such as this convention provide valuable opportunities for young minds to exchange ideas, explore emerging possibilities and contribute to meaningful academic and technological advancements.

I congratulate ISTE Bihar & Jharkhand Chapter and all those associated with this initiative. I wish the convention great success and hope its deliberations inspire fresh perspectives and meaningful collaborations.

Let us not merely prepare our students for the future of technology; let us empower them to shape that future.

With my best wishes.

Prof. (Dr.) D. K. Singh
Vice-Chancellor
Jharkhand University of Technology, Ranchi

बी. आई. टी. सिन्दरी
उच्च एवं तकनीकी शिक्षा विभाग
झारखण्ड सरकार



B.I.T. SINDRI
Department of Higher and Technical Education
Govt. of Jharkhand
P.O. - Sindri Institute, Dhanbad-828123 (Jharkhand)
Ph.-0326-2350495 (O) 0326-2350496 (R)
Email id.: director@bitsindri.ac.in

*It gives me enormous pleasure to express my greetings and best wishes on the moment to all the distinguished dignitaries, academicians, industry professionals, faculty members, and student participants attending "**3rd ISTE BIHAR & JHARKHAND STUDENT'S CONVENTION"** on "**Transformational changes in technical education through Artificial Intelligence"** organized by "**Guru Gobind Singh Educational Society's Technical Campus (An Engineering & Management College**" during **Nov 18-19, 2026** at **Kandra (Vill), Chas, Bokaro, Jharkhand – 827013 .***

In the era of rapid technological progression and economic revolution the theme of this convention, "**Transformational Changes in Technical Education through Artificial Intelligence**" is highly relevant for skill development and the creation of robust industrial and technological ecosystems are essential for building a self-reliant, inclusive, and progressive nation. I believe that Artificial Intelligence is no longer merely an emerging technology; it has become a powerful catalyst for transforming the way we learn, teach, innovate, and solve real-world problems in nurturing students, encouraging innovation, and strengthening collaboration among academia, industry, entrepreneurs, and policymakers. **Technical education** stands at the threshold of a new era where **AI** is redesigning conventional teaching-learning methodologies, enabling adapted learning, supporting insightful assessment, intensifying research capabilities, and promoting innovation. The convergence of skill development, technical education and AI provide an excellent platform for young minds to exchange ideas, showcase their talents, interact with experts, and develop a collaborative spirit. Academic institutions like BIT Sindri remain committed to promoting quality technical education, research, innovation, and industry-academia collaboration to prepare our students for the challenges and opportunities of the future.

I appreciate the organizers, particularly the ISTE Bihar & Jharkhand Section and the student organizing team, for taking this commendable initiative and hope that the deliberations and interactions during the event will generate constructive ideas, actionable strategies, and lasting partnerships in these vital areas. I extend my sincere congratulations to the organizing team and my best wishes for the grand success of **3rd ISTE BIHAR & JHARKHAND STUDENT'S CONVENTION**.

With warm regards and best wishes,

B.I.T. SINDRI Dhanbad (Jharkhand)

Prof. (Dr.) Pankaj Rai,
Director, BIT Sindri

Message from the President



It gives me immense pleasure and is a matter of great privilege for us to be associated with the 3rd ISTE Bihar & Jharkhand Student Convention 2026, centred on the theme *"Transformational Changes in Technical Education through Artificial Intelligence."*

On this significant occasion, I extend my warm greetings and best wishes to the organisers, distinguished dignitaries, faculty members, participating institutions and, most importantly, the student fraternity.

Artificial Intelligence is reshaping the landscape of technical education, opening new avenues for innovation, research, creativity and interdisciplinary learning. As we embrace this transformative era, it is imperative that our students not only equip themselves with emerging technologies but also develop critical thinking, ethical awareness and the ability to use technology responsibly for the greater good of society.

I firmly believe that platforms such as this convention play a vital role in nurturing young minds, encouraging meaningful exchange of ideas and preparing future professionals to meet the challenges and opportunities of a rapidly evolving technological world.

I congratulate the Society for Information Technology in Education (SITE), Bihar & Jharkhand Chapter, and all those associated with organising this convention. I wish the event great success and hope that its deliberations inspire fresh perspectives, meaningful collaborations and innovative ideas for the future of technical education.

With my warm regards and best wishes.

Shri Tarsem Singh

President
GGES, Chief Patron

Message from Prof. (Dr.) Anil Kumar, FIE



Dear Students, Faculty Members, and Distinguished Guests,

It gives me immense joy and a deep sense of pride to send my warmest greetings and best wishes for the 3rd ISTE Bihar & Jharkhand Students' Convention, being held on 18th and 19th November 2026 at the Guru Gobind Singh Educational Society's Technical Campus, Bokaro.

The theme you have chosen - ***"Transformational Changes in Technical Education through Artificial Intelligence"*** - could not be more timely. We live in an era where AI is not simply a subject to be studied; it is reshaping the very nature of how we learn, how we work, and how we solve the problems that matter. For students of engineering and technology, this is not a distant future to prepare for — it is the present you are already stepping into.

I wish to express my sincere and heartfelt appreciation to Guru Gobind Singh Educational Society's Technical Campus for taking this initiative, and in particular to its Director, Prof. (Dr.) Priyadarshi Jaruhar, whose leadership, vision and tireless commitment to quality technical education have been an inspiration to all of us in this region. Prof. Jaruhar exemplifies what is best in our profession — a genuine educator who places students and institutional integrity at the centre of everything he does. Under his stewardship, GGSESTC has earned the respect of the entire ISTE Bihar–Jharkhand community, and this Convention is yet another proof of that spirit.

To every student participating in this Convention - online or in person - I have one message: this gathering is not just an academic event. It is an opportunity to think bigger, ask harder questions, and connect with fellow students and teachers who share your passion for learning and for the future of this country.

Technical education in India is at an inflection point. The decisions taken and the habits of mind formed in institutions like yours today will determine whether India truly leads the world in engineering, innovation and sustainable development in the decades ahead. That is a responsibility worth embracing with excitement, not anxiety.

I send my very best wishes to every student, every faculty member, and everyone associated with this Convention. May it be a memorable and enriching experience for all.

With warm regards and high hopes for all of you,

Prof. (Dr.) Anil Kumar, FIE
National Executive Council Member, ISTE
Ex-Officio Member, ISTE Bihar–Jharkhand Section Managing Committee
Professor of Practice, NITTTR Kolkata
(Deemed University, Ministry of Education, GoI)
Former: DDG & SDG I/C, IEI | Senior Advisor, BSDMA, GoB | State Head, TEQIP-III, MoE
Mobile: +91 9007560314 | kranil01011962@gmail.com

Message from the Secretary



I take this opportunity to convey my sincere appreciation to everyone associated with The **3rd ISTE Bihar & Jharkhand Student Convention 2026**, hosted by Guru Gobind Singh Educational Society's Technical Campus, Bokaro. This marks a meaningful initiative towards encouraging student participation, academic exchange and technological awareness. I am pleased to be associated with this important gathering, centred on the theme *"Transformational Changes in Technical Education through Artificial Intelligence,"* and extend my best wishes to everyone involved.

I strongly believe that the true measure of an educational institution lies not only in the knowledge it imparts but also in the opportunities it creates for its students. Platforms such as the SITE Student Convention provide young learners with a valuable opportunity to showcase their ideas, share their perspectives, interact with fellow students and academicians, and develop the confidence to participate in meaningful academic and technological discussions.

At GGSESTC, we have always believed in encouraging students to step beyond the boundaries of their classrooms and actively engage in activities that contribute to their intellectual and professional growth. The enthusiastic participation of our students and the dedicated efforts of our faculty members and organising team in making this convention possible reflect the institution's collective commitment to creating such enriching experiences. It is their involvement, teamwork and perseverance that bring an academic initiative of this scale to life.

The theme of this convention is particularly relevant in the present times. Artificial Intelligence is steadily transforming the way education is delivered, accessed and experienced. It has the potential to make learning more personalised, accessible, innovative and interdisciplinary. However, the future of education must not be defined by technology alone. The ability to think independently, question intelligently, exercise sound judgement and use technology ethically will remain equally important. Our students must learn to work with AI, while retaining the human creativity and critical thinking that give knowledge its true value.

I appreciate ISTE Bihar & Jharkhand Chapter for providing students with a platform that encourages participation, collaboration and the exchange of ideas. I also congratulate the entire GGSESTC fraternity for their sincere efforts in hosting this convention.

May this gathering inspire our students to become confident thinkers, creative innovators and responsible contributors to society. I wish all participants an enriching experience and the convention great success.

Shri Surendra Pal Singh
Secretary
Guru Gobind Singh Educational Society, Bokaro

Message from the Director



It gives me immense pleasure to extend my heartfelt greetings and best wishes on the occasion of the **3rd SITE Bihar & Jharkhand Student Convention 2026**, hosted by Guru Gobind Singh Educational Society's Technical Campus, Bokaro, on the theme *"Transformational Changes in Technical Education through Artificial Intelligence."*

The world of technical education is undergoing a profound transformation. Artificial Intelligence, emerging technologies and digital innovation are redefining the way knowledge is created, disseminated and applied. As educational institutions, our responsibility extends beyond imparting technical knowledge; we must nurture inquisitive minds, encourage research and innovation, and prepare students to become responsible professionals capable of addressing the challenges of tomorrow.

At GGSESTC, our journey since 2011 has been guided by a commitment to academic excellence, holistic development and meaningful contributions to society. Over the years, the institution has steadily expanded its academic portfolio across Engineering, Management and Computer Applications, including the introduction of Diploma programmes in Civil, Mechanical and Electrical & Electronics Engineering, along with opportunities for working professionals to pursue B.Tech programmes. Our achievement of NAAC B++ accreditation in 2025 marks an important milestone in this continuing journey of institutional development.

We have consistently worked towards strengthening our academic ecosystem through the implementation of NEP 2020 initiatives, Outcome-Based Education, ICT-enabled teaching-learning practices, MOOCs and NPTEL, faculty development programmes and industry-oriented learning. Our collaborations with organisations such as Vedanta, Adani Cement, Powerica and Medicant Hospital, along with internship and industrial training opportunities and engagement with over 45 recruiting organisations, reflect our endeavour to bridge the gap between classroom learning and professional practice.

Our commitment to research, innovation and professional development has also received recognition through various academic initiatives and achievements, including the ISTE Best Faculty Chapter Award, AICTE-supported ATAL Faculty Development Programmes, national conferences and the Jharkhand Science and Innovative Technology Exhibition. We continue to encourage students to participate in research, technical competitions, entrepreneurship initiatives and community-oriented activities, recognising that education must contribute to both individual growth and societal progress.

Hosting this convention is an opportunity to bring together academicians, researchers, industry representatives and young learners to exchange ideas and explore the possibilities that Artificial Intelligence offers to technical education. I particularly encourage students to approach AI not merely

as a tool of convenience but as an instrument of creativity, critical thinking, innovation and responsible problem-solving.

I sincerely appreciate the efforts of ISTE Bihar & Jharkhand Chapter and the organising team for bringing this important academic gathering to our institution. I extend my best wishes to all the delegates and participants and hope that the convention generates meaningful discussions, new perspectives and lasting academic collaborations.

At GGSESTC, we believe that the future of education belongs to those who combine technological excellence with human values, innovation with responsibility, and knowledge with the courage to create change.

I wish the convention grand success.

Dr. Priyadarshi Jaruhar
Director
Guru Gobind Singh Educational Society's Technical Campus
Kandra, Bokaro, Jharkhand

Table of Content

S.N.	Title	Page No.
A	Copyright page	i
B	Convention Organization	ii-iii
C	Messages	iv-xi
D	Table of Content	xii-xv
1.	Artificial Intelligence and Digital Innovation for Social Value Creation: An Empirical Study of Small Retail Businesses in Bokaro, Jharkhand **Samaya Fatima, Dr. Shafaque Rahmat**	1
2.	Beyond Human Hands: Impact of AI Tools on Business Marketing Efficiency **Sneha Rani, Aman Raj, Shafaque Zaheer**	7
3.	Digital Innovation in Small Social Enterprises in Jharkhand **Maryam Shabbir**	12
4.	Can Artificial Intelligence Replace Human Managers? A Study of Opportunities and Limitations **Rani Murmu, Shahnaaz Farhin**	19
5.	E-Learning among Undergraduates: Prospects and Challenges—A Case Study of East Singhbhum, Jharkhand **Sania Perween, Uzma Ali**	23
6.	EduBhasha-AI: Intelligent Bilingual Pedagogy and Error Resolution Framework for Regional Technical Education **Ashu Kumar**	29
7.	Mining and rural livelihoods in jharkhand: assessing the socio-economic and environmental impacts on sustainable development **Harisha Kumari, Saloni Kumari, Shafaque Zaheer**	34
8.	Impact of digital transformation on employee engagement, work–life balance and organisational resilience **Aman Aftab, Adnan Aslam, Shafaque Rahmat**	40
9.	Beyond Credit: Digital Financial Inclusion and AI Readiness among Women in Self Help Groups in Jharkhand **Uzma Aliya**	47
10.	Waste to Wealth: Industrial Waste Reuse in Selected Manufacturing Units of Adityapur Industrial Area, Jharkhand **Sania Salim**	52
11.	An Intelligent Edge-AI Framework for Low-Latency and Energy-Efficient Internet of Things Services **Chandan Kumar, Sneha Kumari, Archita Kumari, Amit Paul**	58

12.	Beyond code correction: a progressive hint-based ai tutoring framework for developing independent debugging and problem-solving skills across programming languages in technical education **Janvi Singh, Anjali Kumari, Amit Paul**	63
13.	Impact of Artificial Intelligence on Digital Marketing for Small Businesses **Shourya Pratap, Amit Paul**	70
14.	InterviewLM: A shareable, retrieval-grounded conversational resume and mock interview platform **Harsh Kumar, Sonu Kumar, Amit Paul**	75
15.	A Pixel-Level Framework for Studying Image Reconstruction Disruption and Protection against Unauthorized Manipulation **Shaikh Affiaa, Amit Paul**	80
16.	Detecting Prompt Injection and Unauthorized Tool Use in AI Agents **Arif Iqbal, Akshat Kr. Singh, Jyoti Kumari, Amit Paul**	85
17.	Applications of Silver Nanoparticles in Modern Science and Technology **Suman Kumari, Janasheen Rani and Deepak Kumar**	89
18.	A Comprehensive Review: Metal Organic Frameworks for H_2 Storage Applications **Dinesh Kumar Gupta, Deepak Kumar**	92
19.	Bioplastics from Chicken Feathers: A Sustainable Solution **Sumit Sagar, Vishal Kumar, Mrinal Manjhi, Manojit De**	97
20.	A Review: Mass Transit System for Sustainable Smart Cities **Manish Kumar Gupta, Prakash Kumar, Dr. Rajendra Prasad Verma**	100
21.	Non-destructive evaluation of an existing reinforced concrete structure using rebound hammer and ultrasonic pulse velocity tests **Sandhya Kumari, Sachin Kumar, Yasir Karim**	104
22.	A Confidence-Gated Naive Bayes Recommender for Data-Driven Crop Selection from Soil Nutrient and Weather Data **Md Shakir Ali, Md Sitare, Tousif Tarique, Majid Alam, Vivek Paramanik**	110
23.	Decoding Autonomy: Monitoring and Evaluating Intelligent AI Agents **Siddhi Choudhary, Shambhavi Mishra, Suman Kumari**	121
24.	AI-based digital payment fraud detection and prevention system **Lilawati Kumari, Kumkum Kumari, Neha Kumari**	125
25.	Artificial Intelligence and Security: Threats, Defenses, and Governance **Sourabh Kumar, Ayush Singh, Md Hussain Ansari**	129

26.	Predictive AI for Disaster Risk Forecasting and Early Warning Systems **Harshit Kriti Raj Mandal, Gourav Gupta, Mrs. Apurba Sinha**	134
27.	Machine Learning-Based Detection of Phishing Websites for Protecting User Data Privacy: A Calibrated and Interpretable Random-Forest Framework **Aariz Abdullah, Faiza Afreen, David Kumar Pandey, Rani Perween, K Shreya Rao**	140
28.	Quantum-Inspired Neural Network for Digital Signature **Aman Kaushik, Kajal Kumari. Arun Prasad Burnwal**	151
29.	Quantifying Resampling-Induced Data Leakage in Imbalanced Credit Card Fraud Detection: A Comparative Study of Sampling Strategies and Classification Models **Ashutosh Prasad, Nitesh Kumar, Smita Kishore**	157
30.	WISEBUD: "An Offline-First AI-Powered Learning and Student Collaboration Platform" **Shreya Gupta, Shreya Anand, Aakanksha Sinha**	164
31.	Changes In Technical Education Using Artificial Intelligence: A Review of Emerging Practices, Opportunities and Challenges **Ashish Raj and Salim Ahmad**	168
32.	A Critical Review of Recently Proposed Metaheuristic Algorithms: Taxonomy, Benchmark Performance, and Open Research Gaps **Prince Kumar, Rohit Verma, and Manoj Kumar**	172
33.	AI driven smart suraksha helmet with edge AI for real time hazard detection **Bhola Kumar Mahto; Sahin Parween and Abhilash Ghosh**	179
34.	Intelligent Control and Automation: Machine Learning, Robotics, and Learning-Enabled Control Systems **Siya Kumari, Neha Kumari, Alok Kumar**	185
35.	Optimization Based Approache for Congestion Management in Power System Transmission Lines **Tanushri Kumari, Neha Kumari Paul, Uttama Kumar Das**	190
36.	Micro-Ring Resonator Based Optical Switch: Principles, Design, Architectures and Challenges **Vikash Gope Alok Kumar and Pankaj Kumar Ray**	193
37.	Green finance in india: emerging opportunities and future prospects **Aarushi Raj, Md Asif Faizi**	199
38.	Impact of artificial intelligence on employee productivity and job opportunities **Muskan Kumari, Muskan Perween, Md Asif Faizi**	205
39.	Adoption of Digital Supply Chain Management and Its Effect on Operational Efficiency in the Indian Automobile Industry **Shalu Kumari, Mahavir Prasad**	211
40.	Supply Chain Adaptiveness and Marketing Responsiveness in the FMCG Sector **Rajnish Kumar, Vikash Kumar Jain**	216
41.	Enhanced Pneumatic Control System for Precise, Jerk-Free Actuation Under Bi-Directional And Vibrating Load Conditions **Akhilesh Anand, Atul Kumar, Kumar Ashish**	223

42.	Variation of Brake Thermal Efficiency of I C Engine with increase in Ethanol in Ethanol-Petrol Blends **Amit Kumar, Gulam Hussin, Dr. Bikash Ghoshal**	227
43.	Comprehensive Review of the Use of Artificial Intelligence in Modern Refrigeration Systems **Khushboo Kumari , Sumit Kumar Pandey**	234
44.	AI Powered vahicle mounted in takication, detection and driver saferty system Khstar Jahan, Iffat Perween, Abhilash Ghosh	243
45.	Pavexa – A Porous Concrete Road Solution for Storm Water Management **Satyam Vishwakarma, Zarnaab Khan, Sanjay Hembram**	248
46.	Dynamic Speed Control of a 3-Phase Squirrel-Cage Induction Motor via Field-Oriented Strategy **Neha Kumari Paul, Tanushree, Jayanti**	251
47.	The role of generative AI in enhancing employee productivity and organizational innovation **Sahiba Akram, Dr. Shafaque Rahmat, Dr. Priyadarshi Jaruhar**	258

Artificial Intelligence and Digital Innovation for Social Value Creation: An Empirical Study of Small Retail Businesses in Bokaro, Jharkhand

Samaya Fatima, BBA, GGSESTC Chas Bokaro
Dr. Shafaque Rahmat, Assistant Professor, MBA Department, GGSESTC Chas, Bokaro

Abstract

Artificial Intelligence (AI) and digital technologies are transforming the operations of small retail businesses. This study examines the influence of AI and digital innovation on social value creation among small retail businesses in Bokaro, Jharkhand. The study focuses on technological, organisational, environmental and human resource factors influencing digital adoption. It further investigates how digital innovation capability contributes to employment development, customer inclusion, and community linkages and workplace well-being. The study adopts a quantitative research design using a structured questionnaire. For demonstration purposes, simulated data representing 120 small retail businesses are analysed using descriptive statistics and Partial Least Squares Structural Equation Modelling (PLS-SEM). The illustrative results indicate that technological factors, organisational support and HR digital capability positively influence technology adoption. Digital adoption also positively influences innovation capability, which contributes to social value creation. The study highlights the importance of digital skills development and employee training in maximising the social benefits of digital transformation.

Keywords: Artificial Intelligence, Digital Innovation, Social Value Creation, Small Retail Businesses, Digital Competence, Bokaro.

1. Introduction

Digital transformation has become an important factor in improving business performance and competitiveness. Artificial Intelligence, digital payments, online marketing, automated billing and inventory management are increasingly being adopted by small retail businesses. In India, small retailers play an important role in employment generation, customer services and local economic development. Bokaro, Jharkhand, provides an interesting context for examining digital transformation because of its industrial development and growing commercial activities.

Although digital technologies can improve business efficiency, their contribution extends beyond financial performance. They can improve employee skills, customer accessibility, supplier relationships and workplace conditions.

However, the successful implementation of digital technology depends on organisational readiness and the digital competence of owners and employees.

Therefore, this study examines how AI and digital innovation contribute to social value creation among small retail businesses in Bokaro.

2. Review of Literature

The Technology–Organisation–Environment (TOE) framework developed by Tornatzky and Fleischer (1990) explains technology adoption through technological characteristics, organisational readiness and environmental influences.

Barney (1991), through the Resource-Based View, emphasises that organisational resources and capabilities can create sustainable competitive advantages.

Davis (1989) explains that perceived usefulness and ease of use influence technology acceptance.

Porter and Kramer (2011) introduced the concept of creating shared value, highlighting the relationship between business activities and societal benefits.

Wright, Dunford and Snell (2001) emphasise the strategic importance of human resources and employee capabilities in organisational performance.

Based on these perspectives, the present study integrates technology adoption, digital innovation capability and human resource competence to examine social value creation in small retail businesses.

3. Research Objectives

1. To examine the extent of AI and digital technology adoption among small retailers in Bokaro.
2. To identify technological, organisational and environmental factors influencing digital adoption.
3. To analyse the relationship between digital adoption and innovation capability.
4. To examine the impact of digital innovation capability on social value creation.
5. To investigate the role of digital competence in strengthening these relationships.

4. Research Hypotheses

- H1: Technological factors positively influence digital adoption.
- H2: Organisational factors positively influence digital adoption.
- H3: Environmental factors positively influence digital adoption.
- H4: HR digital capability positively influences digital adoption.
- H5: Digital adoption positively influences digital innovation capability.
- H6: Digital innovation capability positively influences social value creation.
- H7: Digital innovation capability mediates the relationship between digital adoption and social value creation.
- H8: Digital competence positively moderates the relationship between adoption and innovation capability.
- H9: Digital competence positively moderates the relationship between innovation capability and social value creation.

5. Research Methodology

The study follows a quantitative, descriptive and explanatory research design.

Study area: Bokaro, Jharkhand.

Target population: Small retail businesses operating in Bokaro.

Sample size: 120 retail businesses (simulated dataset).

Sampling technique: Stratified sampling (proposed).

Data collection instrument: Structured questionnaire.

Measurement scale: Five-point Likert scale, ranging from 1 (Strongly Disagree) to 5 (Strongly Agree).

Statistical techniques: Percentage analysis, descriptive statistics, reliability analysis and PLS-SEM.

6. Data Analysis and Interpretation

6.1 Demographic Profile of Respondents

The simulated sample consists of 120 small retail businesses.

Retail category	**Respondents**	**Percentage**
Grocery and general stores	40	33.3%
Apparel and footwear	25	20.8%
Electronics and mobile	20	16.7%
Pharmacy and health	15	12.5%
Hardware and others	20	16.7%
Total	**120**	**100%**

Interpretation: Grocery and general stores constitute the largest proportion of the sample (33.3%), followed by apparel and footwear businesses (20.8%). This indicates that the illustrative sample represents different categories of small retail businesses.

6.2 Adoption of Digital Technologies

Digital technology	**Businesses adopting**	**Percentage**
Digital payments (UPI/cards)	110	91.7%
Digital billing/inventory	58	48.3%
Social media and messaging sales	66	55.0%
AI-enabled business tools	22	18.3%

Interpretation: Digital payments represent the highest level of adoption (91.7%). However, only 18.3% of businesses use AI-enabled tools. This suggests that small retailers are more familiar with basic digital technologies than advanced AI applications. Training and awareness programmes may help bridge this gap.

6.3 Descriptive Analysis of Social Value Creation

The following table presents illustrative mean scores on a five-point Likert scale.

Social value dimensions	**Mean**	**SD**
Employment and skill development	3.42	0.71
Customer inclusion and access	3.78	0.68
Community and supplier linkages	3.31	0.74
Workplace well-being	3.25	0.79

Interpretation: Customer inclusion and access have the highest mean score (3.78), indicating that digital payments and online ordering can improve customer accessibility. Workplace well-being records the lowest mean score (3.25), suggesting that digital adoption does not automatically improve employees' working conditions.

6.4 Reliability and Validity Analysis

Illustrative measurement model results are presented below.

Construct	**Cronbach's Alpha**	**CR**	**AVE**
Technology factors	0.82	0.87	0.58

Organisation factors	0.84	0.89	0.62
Environment factors	0.79	0.85	0.54
HR digital capability	0.86	0.90	0.60
Adoption intensity	0.83	0.88	0.59
Digital innovation capability	0.85	0.89	0.62
Social value creation	0.88	0.91	0.60

Interpretation: All illustrative Cronbach's Alpha and Composite Reliability values exceed 0.70, while AVE values exceed 0.50. This indicates satisfactory internal consistency and convergent validity in the simulated example. Actual reliability and validity must be established using the collected survey responses.

6.5 Structural Model and Hypothesis Testing

The following simulated PLS-SEM results demonstrate the proposed relationships.

Hypothesis	**Relationship**	**β**	**p-value**	**Decision**
H1	Technology → Adoption	0.24	0.006	Supported
H2	Organisation → Adoption	0.18	0.028	Supported
H3	Environment → Adoption	0.14	0.071	Not supported
H4	HR capability → Adoption	0.31	0.001	Supported
H5	Adoption → Innovation capability	0.62	0.001	Supported
H6	Innovation capability → Social value	0.58	0.001	Supported
H7	Adoption → Innovation → Social value	0.36	0.001	Supported
H8	Adoption × Competence → Innovation	0.16	0.018	Supported
H9	Innovation × Competence → Social value	0.12	0.047	Supported

Note: p-values of 0.001 represent $p < 0.001$.

Interpretation:

The simulated results show that HR digital capability has the strongest influence on technology adoption ($\beta = 0.31$), followed by technological factors ($\beta = 0.24$).

Environmental factors have a positive but statistically insignificant relationship with adoption ($p = 0.071$).

Digital adoption strongly influences digital innovation capability ($\beta = 0.62$). Furthermore, innovation capability positively influences social value creation ($\beta = 0.58$).

The mediation result ($\beta = 0.36$) suggests that digital innovation capability acts as a mechanism through which digital adoption contributes to social value creation.

The positive moderation coefficients indicate that digital competence strengthens both relationships.

6.6 Coefficient of Determination (R^2)

Endogenous construct	**R^2**	**Explained variance**

Digital adoption	0.58	58%
Digital innovation capability	0.42	42%
Social value creation	0.34	34%

Interpretation: The model explains 58% of the variance in digital adoption, 42% in digital innovation capability and 34% in social value creation. These illustrative values indicate reasonable explanatory potential for the proposed research model.

7. Major Findings

Based on the illustrative analysis, the following findings emerge:

1. Digital payments are the most widely adopted technology among small retailers.
2. AI-enabled tools have relatively low adoption.
3. HR digital capability is an important predictor of digital adoption.
4. Digital adoption positively influences innovation capability.
5. Digital innovation capability contributes to social value creation.
6. Customer inclusion is the strongest-performing social value dimension.
7. Digital competence strengthens the relationship between technology adoption, innovation and social value creation.

8. Managerial and Practical Implications

The study suggests several practical implications:

Digital Skill Development: Small retailers should provide regular digital training to employees and owners.

Employee Development: Businesses should encourage employees to learn digital billing, inventory management and AI-enabled applications.

Customer Inclusion: Digital payments and online ordering should be used to improve accessibility for different customer groups.

Government Support: Local authorities and industry associations should organise digital awareness and AI training programmes for small retailers.

Technology Providers: Affordable, simple and regional-language digital solutions should be developed for small businesses.

9. Conclusion

The study examines the relationship between artificial intelligence, digital innovation and social value creation among small retail businesses in Bokaro, Jharkhand.

The illustrative analysis indicates that technology adoption alone may not be sufficient to generate social benefits. Human resource capability, employee digital competence and innovation practices are important in converting technology adoption into meaningful outcomes.

The study contributes to the existing literature by integrating the TOE framework, Resource-Based View and strategic human resource management perspectives.

It concludes that digital transformation should not be considered only a means of improving business efficiency but also an opportunity to promote employment development, customer inclusion, community relationships and workplace well-being.

10. Limitations and Future Research

The study is limited to small retail businesses operating in Bokaro, Jharkhand, which may restrict the generalisability of its findings to other geographical regions. The cross-sectional research design also limits the ability to establish long-term causal relationships among the variables. Future research may extend the study to other districts of Jharkhand and compare urban and rural retail businesses. Longitudinal studies may further examine how AI adoption and digital innovation influence social value creation over time.

References

1. Austin, J., Stevenson, H., & Wei-Skillern, J. (2006). Social and commercial entrepreneurship: Same, different, or both? *Entrepreneurship Theory and Practice, 30*(1), 1–22.
2. Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management, 17*(1), 99–120.
3. Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319–340.
4. Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review, 31*(1), 2–24.
5. Porter, M. E., & Kramer, M. R. (2011). Creating shared value. *Harvard Business Review, 89*(1–2), 62–77.
6. Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). Free Press.
7. Tornatzky, L. G., & Fleischer, M. (1990). *The processes of technological innovation*. Lexington Books.
8. Wright, P. M., Dunford, B. B., & Snell, S. A. (2001). Human resources and the resource-based view of the firm. *Journal of Management, 27*(6), 701–721.

Beyond Human Hands: Impact of AI Tools on Business Marketing Efficiency

Sneha Rani
Department of Bachelor of Business Administration,
Guru Gobind Singh Educational Society's Technical Campus, Bokaro India
sr0919078@gmail.com

Aman Raj
Department of Bachelor of Business Administration,
Guru Gobind Singh Educational Society's Technical Campus, Bokaro India
amanraj246452@gmail.com

Shafaque Zaheer
Department of Bachelor of Business Administration,
Guru Gobind Singh Educational Society's Technical Campus, Bokaro India
shafaque.zaheer@gmail.com

ABSTRACT
This study examines the impact of AI tools on the efficiency of business marketing. It is based on secondary research, reviewing recent industry surveys and reports on the use of AI in marketing. A study by SAS and Coleman Parkes, covering 300 organisations worldwide, found that 85% of marketers were actively using generative AI in 2025, up from 75% in 2024. A Salesforce and YouGov survey of 1,029 marketers found that they expect generative AI to save about five hours of work every week, with writing marketing copy and basic content creation among the most common uses for current users. The same survey also highlighted concerns about the quality and accuracy of AI outputs, as well as a lack of skills to use these tools effectively. Similarly, an Econsultancy survey of over a thousand marketers found that only 26% described their organisation's digital skills development as fit for purpose. Overall, the review suggests that marketers expect AI tools to improve efficiency by saving time and speeding up content creation. However, accuracy checks and proper training remain necessary. This study helps businesses and students understand where AI can support marketing and where human judgement is still required.

Keywords: Artificial Intelligence, AI Tools, Business Marketing, Marketing Efficiency, Generative AI, Content Creation, Digital Skills, Human Judgement.

1. INTRODUCTION

Marketing has long depended on people to write copy, prepare campaigns, study customer data and answer enquiries. Generative artificial intelligence (GenAI) has begun to change this. Chatbots, large language models and image generators can now draft text, summarise data, suggest campaign ideas and respond to customers at a speed that manual work cannot match. In this paper, marketing efficiency is understood simply as the amount of useful marketing output (content, campaigns, customer responses) produced for a given amount of time, cost and effort. Industry surveys report fast adoption and large expected benefits, but two questions remain for managers and students. First, does wider use of AI tools actually translate into efficiency gains for marketing teams? Second, which marketing tasks can be handed to AI, and which still need a person's judgement? This paper addresses both questions by bringing together findings from three recent industry surveys and from experimental research on generative AI at work. The title, "Beyond Human Hands", reflects the central idea: AI can take over much of the routine effort, yet human oversight does not disappear.

The study has three objectives: (i) to examine how widely AI tools are used in business marketing; (ii) to assess the efficiency benefits that marketers report or expect, especially time saving and content creation; and (iii) to identify the challenges (output quality, accuracy, skills and governance) and the areas where human judgement is still required.

2. LITERATURE REVIEW AND THEORETICAL FRAMEWORK

2.1 AI and the future of marketing

Davenport et al. (2020) proposed a framework for understanding how AI will influence marketing strategies and customer behaviour, and argued that AI is more effective when it augments human managers rather than replaces them. Their view frames the central tension of this paper: AI can raise efficiency, but it works best alongside human decision-makers.

2.2 Evidence on productivity

Controlled studies provide evidence beyond opinion surveys. In a preregistered experiment with 453 college-educated professionals, Noy and Zhang (2023) found that ChatGPT reduced the average time taken for professional writing tasks by 40% and raised output quality by 18%. Brynjolfsson et al. (2025) studied 5,179 customer support agents and found that access to a generative AI assistant raised the number of issues resolved per hour by about 14% on average, with much larger gains for novice and low-skilled workers. Both findings are relevant to marketing, where writing and customer interaction are core activities.

2.3 Limits of AI and the need for judgement

Dell'Acqua et al. (2023) experimented with 758 consulting professionals and described a "jagged technological frontier". For tasks inside the frontier of AI capability, consultants using AI were more productive and produced better work. For a task deliberately placed outside it, however, those using AI were 19 percentage points less likely to reach a correct solution. This result supports the need for accuracy checks and trained users, which is also the practical message of the surveys reviewed below.

The Technology Acceptance Model (Davis, 1989) explains technology use through perceived usefulness and perceived ease of use. In this paper, marketers' expectations of time saving and faster content creation represent perceived usefulness, while the reported lack of skills, training and governance represents barriers to ease of use. The findings in Section 5 are organised around this logic.

3. RESEARCH METHODOLOGY

The study follows a descriptive, comparative design based entirely on secondary data. Three industry surveys on marketers' use of AI were selected because they (a) focus on marketing professionals, (b) report clear sample sizes and fieldwork periods, and (c) were published by identifiable organisations between 2023 and 2025. They are complemented by three peer-reviewed or working-paper studies on generative AI and productivity, identified through Google Scholar searches on generative AI, marketing efficiency and digital skills. Percentages are reproduced as published by the survey organisations and grouped into four themes: adoption, efficiency, use cases, and barriers. Microdata was not available, so no statistical testing was attempted. Table 1 summarises the survey sources.

Table 1. Industry surveys reviewed

Source	Timing	Sample	Indicators used in this paper
SAS and Coleman Parkes, Marketers and AI: Navigating New Depths	2025 (fieldwork June–July); compared with a 2024 wave	300 marketing decision-makers, worldwide	GenAI adoption, understanding, top uses, benefits, ROI, governance
Salesforce and YouGov, Generative AI Snapshot	May 2023	1,029 full-time marketers (US, UK, Australia)	Current and planned use, expected weekly time saving, content creation, skills and training gaps
Econsultancy, The Future of Marketing 2025	2025 report	1,096 brand, vendor and agency marketers	Digital skills development fit for purpose

4. DATA ANALYSIS AND FINDINGS

4.1 Adoption Is Widespread and Rising

The SAS and Coleman Parkes study reports that 85% of marketers were actively using GenAI in 2025, compared with 75% in 2024, while the share reporting a strong understanding of GenAI and its business impact rose from 50% to 62%. In 2025, 93% of marketing teams were budgeting for GenAI through 2026. The earlier Salesforce and YouGov survey (2023) found 51% already using or experimenting with GenAI and a further 22% planning to start soon, together nearly three-quarters of respondents. Because the samples, regions and definitions differ, the two surveys are not strictly comparable, but both point in the same direction: adoption has moved from early experimentation to routine use (Figure 1a).

4.2 Efficiency Benefits Reported and Expected

In the SAS study, 94% of marketers reported better personalisation, 91% reported more efficient processing of large datasets, and 90% reported savings in time and operating cost. Return on investment was reported by 83% of marketing teams and 93% of chief marketing officers using GenAI. In the Salesforce and YouGov survey, marketers estimated that GenAI would save them more than five hours of work per week, and 71% expected it to remove busy work. It should be noted that this figure is an expectation, not a measured saving. The experimental studies in Section 3.2 supply measured evidence of a similar kind: a 40% reduction in writing time (Noy & Zhang, 2023) and about 14% more customer issues resolved per hour (Brynjolfsson et al., 2025).

4.3 Content Creation is the Leading Use

Content-related work is the most visible application. In the Salesforce and YouGov survey, 76% of marketers using GenAI applied it to content creation and writing copy. In the SAS study, text and copy generation use rose from 34% to 45% between 2024 and 2025, and chatbot and customer-interaction use rose from 43% to 62%. Trend analysis (36%) was the third most common use.

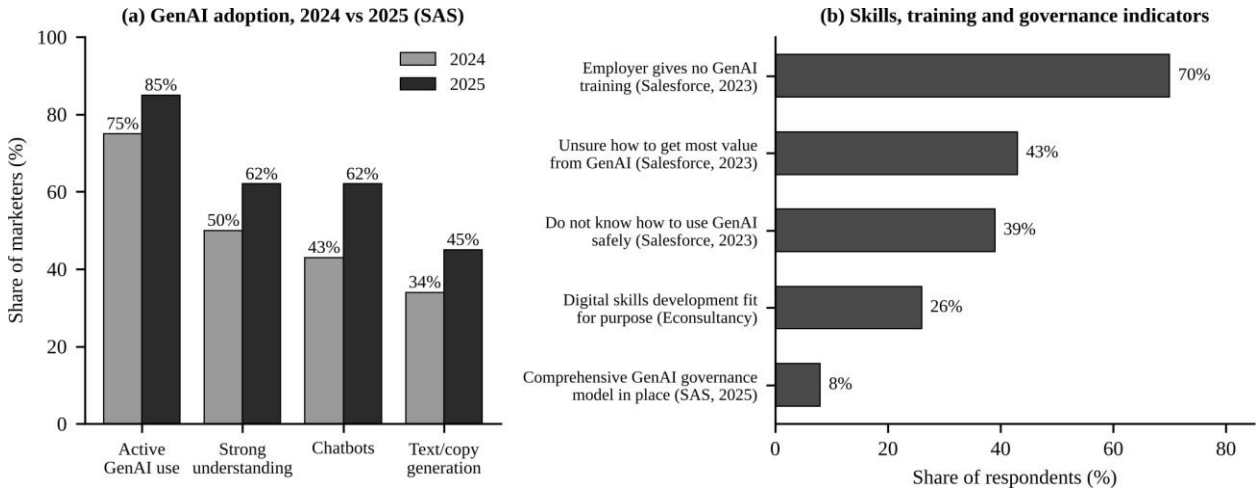


Figure 1. (a) Growth in GenAI adoption indicators among marketers, 2024 to 2025 (SAS and Coleman Parkes). (b) Skills, training and governance indicators from three different surveys. For the first three bars a higher value means a larger gap; for the last two a lower value means a larger gap. The surveys differ in year and sample and are shown side by side for illustration only.

4.4 Barriers: Quality, Skills and Governance

The same surveys show clear limits to efficiency. Marketers in the Salesforce and YouGov study raised concerns about the accuracy and quality of GenAI outputs. In the same survey, 39% said they did not know how to use GenAI safely, 43% did not know how to get the most value from it, and 70% said their employer did not yet provide GenAI training. The Econsultancy survey shows that the problem is broader than AI: only 26% of marketers described their organisation's digital skills development as fit for purpose. On governance, the SAS study found that only 8% of organisations had a comprehensive GenAI governance model in 2025, against 7% in 2024 (Figure 1b).

4.5 Where Human Judgement Is Still Required

Taken together, the surveys and experiments indicate that AI is most useful for routine, high-volume and clearly defined tasks, such as first drafts of copy, chatbot replies and trend summaries. Judgement is still needed to check facts and brand fit, to decide which AI suggestions to use, and to recognise when a task falls outside what the tool can do reliably, as the "jagged frontier" result shows. This agrees with the view of Davenport et al. (2020) that AI should augment, not replace, human managers.

5. DISCUSSION

The findings indicate that GenAI is steadily becoming a part of everyday marketing practices. As reflected in Graph (a), active usage increased from 75% in 2024 to 85% in 2025. The use of chatbots also rose from 43% to 62%, while text and copy generation increased from 34% to 45%. This suggests that marketers are increasingly relying on AI for content creation, customer interaction and routine communication tasks. However, the rise in adoption does not necessarily mean that users have developed an equally strong understanding of the technology. While 85% reported active usage, only 62% demonstrated a strong understanding of GenAI, indicating a gap between using these tools and knowing how to use them effectively.

Graph (b) brings out another important aspect of this growing dependence on AI. Around 70% of respondents reported that their employers did not provide GenAI training, while 43% were unsure about how to derive maximum value from these tools. The fact that only 8% reported having a comprehensive GenAI governance model further points to the lack of organisational preparedness. These findings suggest that businesses cannot rely on technology alone to improve marketing outcomes. Regular training, human supervision and clear guidelines are equally important. For marketing students and future professionals, this also highlights the need to develop AI literacy alongside creativity, communication skills and critical thinking. Ultimately, the value of GenAI will depend not just on how widely it is adopted, but on how thoughtfully and responsibly it is used.

Overall, the findings indicate that the future of AI-driven marketing will depend not simply on the pace of technological adoption but on the ability of individuals and organisations to develop the knowledge, skills, and governance mechanisms necessary to use these technologies effectively. While the increasing adoption figures reflect growing confidence in GenAI's usefulness, the evident gaps in training and governance demonstrate that meaningful adoption requires a more comprehensive approach. The real competitive advantage is therefore likely to emerge not from access to AI tools alone, but from the ability to combine technological efficiency with human intelligence, critical thinking, creativity, and responsible decision-making.

6. LIMITATIONS AND FUTURE SCOPE

This study relies on secondary data and therefore inherits the limits of its sources. SAS and Salesforce sell analytics and AI products, so their survey framing may favour positive findings. The surveys differ in year, region and definitions, so cross-survey comparisons are indicative only. Expected time savings are self-reported expectations, and figures were taken from the publishers' published summaries rather than from raw data. Because AI tools change quickly, the findings are a snapshot. Future work could collect primary data from Indian small and medium enterprises or measure time saved in an experiment with marketing teams.

7. CONCLUSION

The reviewed evidence suggests that AI tools are now a routine part of business marketing and that marketers expect them to improve efficiency, mainly by saving time and speeding up content creation. Experimental studies support the direction of these expectations. The benefits are not automatic, however: concerns about output quality, gaps in skills and training, and limited governance mean that accuracy checks and proper training remain necessary. AI can lift much of the routine workload, but marketing decisions that depend on judgement, context and responsibility still rest with people.

REFERENCES

[1] Brynjolfsson, E., Li, D., & Raymond, L. R. (2025). Generative AI at work. The Quarterly Journal of Economics, 140(2), 889–942. (Working paper version: NBER Working Paper No. 31161, 2023.)

[2] Davenport, T., Guha, A., Grewal, D., & Bressgott, T. (2020). How artificial intelligence will change the future of marketing. Journal of the Academy of Marketing Science, 48(1), 24–42. https://doi.org/10.1007/s11747-019-00696-0

[3] Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319–340. https://doi.org/10.2307/249008

[4] Dell'Acqua, F., McFowland III, E., Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K. C., Rajendran, S., Krayer, L., Candelon, F., & Lakhani, K. R. (2023). Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality (Harvard Business School Technology & Operations Mgt. Unit Working Paper No. 24-013). https://doi.org/10.2139/ssrn.4573321

[5] Econsultancy. (2025). The future of marketing 2025. https://econsultancy.com/reports/the-future-of-marketing-2025/

[6] Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. Science, 381(6654), 187–192. https://doi.org/10.1126/science.adh2586

[7] Salesforce. (2023). New research: 60% of marketers say generative AI will transform their role, but worry about accuracy [Survey conducted with YouGov, May 18–25, 2023; n = 1,029]. Salesforce News. https://www.salesforce.com/news/?p=66319

[8] SAS Institute, & Coleman Parkes Research. (2025). Marketers and AI: Navigating new depths [Survey of 300 marketing decision-makers; press release dated September 15, 2025]. https://www.prnewswire.com/news-releases/new-study-genai-hype-is-over-as-93-of-cmos-see-strong-roi-302555662.html

Digital Innovation in Small Social Enterprises in Jharkhand

Maryam Shabbir
BBA Semester I
Al Kabir Institute of Management and Technology
Email: Sushmita.choudhury@akp.ac.in

Abstract

This study examined a framework linking digital adoption with organisational capabilities, access conditions and perceived outcomes in small social enterprises, with Jharkhand as its intended setting. It reviewed relevant scholarship and evaluated the reporting structure of an existing analytical file containing 200 simulated enterprise records. Digital adoption was measured through the reported frequency of using eight tools, while seven multi-item constructs represented digital skills, connectivity, affordability, language accessibility, organisational support, business benefits and community service. The supplied results included descriptive statistics, Cronbach's alpha estimates and two-sided Spearman correlations. Mean adoption was 1.94 on a zero-to-four scale, and perception means ranged from 2.90 to 3.00. Reported alpha coefficients ranged from .863 to .913. All seven associations with adoption were positive and statistically significant within the supplied file, with coefficients between .702 and .798. These relationships demonstrated the operation of the framework but did not establish empirical conditions among enterprises in Jharkhand. The analysis highlighted the need to distinguish technological access from practical capability and commercial benefits from community outcomes. It produced a structured basis for subsequent field research and identified training, affordability, language access and assisted service delivery as relevant areas for investigation.

Keywords: digital innovation; social enterprises; Jharkhand; digital adoption; organisational capabilities; social inclusion.

Introduction:

Small social enterprises operated at the intersection of commercial activity and social purpose. Their ability to sustain services depended on earning revenue while remaining responsive to the communities they intended to serve. Digital innovation offered a useful lens for examining this relationship because it connected everyday business practices with questions of access, participation and organisational capacity.

The study approached digital innovation as a change in working practices enabled by digital technology. This understanding included ordinary applications used to manage payments, customer communication, inventory and records. The analytical emphasis was placed on what changed within the enterprise rather than on the novelty of the technology. Frequent use, by itself, was not treated as evidence of improved performance or greater inclusion.

The framework adopted a working definition of a social enterprise as an organisation that combined revenue-generating activity with an explicit social or environmental mission. This definition provided a basis for distinguishing the intended research population from businesses whose social contribution had not been documented. It also allowed business benefits and community service to be examined separately.

Jharkhand supplied the geographical focus of the framework. Agriculture and food activities, artisan production, education and skills, health and wellbeing, and environmental services were retained as relevant activity categories. The study did not establish their prevalence in the state. Its central question concerned how adoption could be examined alongside the skills, resources and support needed to integrate digital tools into routine work.

Review of literature:

Qureshi, Pan and Zheng (2021) connected digital technologies with social innovation through a research framework that directed attention to both technological processes and social aims. Their contribution informed the distinction made in this study between using a digital tool and obtaining a socially meaningful result. The framework

consequently included community service as a separate domain rather than treating business performance as a sufficient indicator of social value.

This concern with context was also reflected in Sengupta et al. (2021), whose analysis of eKutir in India examined the conditional acceptance of a digitised business model from a stakeholder perspective. Their work provided a basis for considering the practical suitability of a tool, including affordability and accessibility. In the present framework, these conditions were represented through separate constructs so that operational difficulties were not reduced to a general measure of willingness to adopt.

Kotiranta et al. (2024) examined digitalisation as a possible driver of social enterprise growth. Their research informed the inclusion of perceived business benefits, covering reach, efficiency, costs and opportunities. The present study nevertheless distinguished these perceptions from verified changes in revenue or profit. This distinction mattered because a favourable survey response did not provide the same evidence as an independently checked financial record.

Organisational arrangements provided another part of the explanation. Hu, Ivanov and Battisti (2026) examined tensions associated with digitalisation and the hybrid character of social enterprises. Their perspective informed attention to training, assistance and the distribution of digital responsibilities. These issues suggested that adoption could involve adjustments to roles and routines as well as the acquisition of devices or applications.

Krombacher et al. (2026) developed a research agenda for digital social innovation that connected technological and social concerns. Read alongside the other contributions, their work supported an approach that considered enabling conditions, organisational practices and outcomes together. The OECD (2023) discussion of the social and solidarity economy further highlighted the diversity of organisational forms, reinforcing the importance of explicit eligibility criteria in local research.

The literature informed the construction of a manageable analytical framework rather than a claim that one questionnaire had already been validated for Jharkhand. The study brought capabilities, access conditions and two outcome domains into a common structure. Its contribution rested on this integration and on careful interpretation of the supplied analysis, rather than on an exhaustive review or an independently established regional research gap.

Research objectives:

The overall objective was to examine a framework for studying digital adoption and its relationships with enabling conditions and perceived outcomes in small social enterprises.

- To describe the digital tools represented in enterprise activities and their frequency of use.
- To examine the relationship between digital skills and digital adoption.
- To assess the relationships of connectivity, affordability, language accessibility and organisational support with adoption.
- To examine the associations of adoption with perceived business benefits and community service.
- To identify implementation challenges and support priorities represented in the supplied analytical file.

Research hypotheses:

Seven pairs of two-sided hypotheses were specified. Each null hypothesis stated that digital adoption had no monotonic association with the relevant construct, while the corresponding alternative stated that an association existed. Spearman's population correlation was denoted by rho. The hypotheses were evaluated within the supplied analytical file; the decisions did not establish population relationships in Jharkhand.

$H_0 1$: Digital adoption had no monotonic association with digital skills (rho = 0).

$H_1 1$: Digital adoption had a monotonic association with digital skills (rho ≠ 0).

$H_0 2$: Digital adoption had no monotonic association with internet connectivity (rho = 0).

$H_1 2$: Digital adoption had a monotonic association with internet connectivity (rho ≠ 0).

$H_0$3: Digital adoption had no monotonic association with affordability (rho = 0).

$H_1$3: Digital adoption had a monotonic association with affordability (rho ≠ 0).

$H_0$4: Digital adoption had no monotonic association with language accessibility (rho = 0).

$H_1$4: Digital adoption had a monotonic association with language accessibility (rho ≠ 0).

$H_0$5: Digital adoption had no monotonic association with organisational support (rho = 0).

$H_1$5: Digital adoption had a monotonic association with organisational support (rho ≠ 0).

$H_0$6: Digital adoption had no monotonic association with perceived business benefits (rho = 0).

$H_1$6: Digital adoption had a monotonic association with perceived business benefits (rho ≠ 0).

$H_0$7: Digital adoption had no monotonic association with perceived community service (rho = 0).

$H_1$7: Digital adoption had a monotonic association with perceived community service (rho ≠ 0).

Research methodology:

The study used a literature-informed framework and a secondary review of the statistical summaries supplied in the manuscript. The analytical material comprised 200 simulated enterprise records, as explicitly documented in the source. No field recruitment, participant interviews or researcher-administered survey was established by the attachment. Respondent-level data and executable analysis files were unavailable, so the reported calculations were retained rather than independently reproduced.

The framework was organised around a cross-sectional enterprise survey supplemented by semi-structured interviews. Purposive selection across localities, activity categories and enterprise sizes was considered suitable for developing a varied field sample, although such selection would not support probability-based population estimates. An explicit mission, revenue-generating activity and a documented size criterion were identified as necessary eligibility conditions. The categories of 1–5, 6–10 and 11–20 workers were treated as analytical groupings rather than statutory classifications.

Digital adoption combined eight tool-frequency measures scored from 0 to 4. Seven perception constructs were represented by three or four items on a five-point agreement scale, with 3 as the neutral midpoint. The full wording of the questionnaire and every response anchor were not recoverable from the supplied summaries. Consequently, the study did not assert that instrument validity or respondent understanding had been established.

Frequencies and percentages described categorical variables. Means and standard deviations summarised scale scores. Cronbach's alpha was used to assess reported internal consistency, and two-sided Spearman correlations were used to examine monotonic relationships. The nominal significance level was .05. Because seven associations were examined, a Bonferroni threshold of approximately .00714 was also applied. Statistical significance was interpreted separately from causation and substantive field relevance.

Data analysis and results:

Rural profiles accounted for 35.5% of the 200 records, while urban and mixed-area profiles accounted for 32.5% and 32.0%, respectively. Worker categories were relatively evenly represented. Artisan products formed the largest activity group at 23.5%, followed by health and wellbeing at 21.0%. These figures described the composition of the analytical file and did not estimate the distribution of enterprises in Jharkhand.

Characteristic	Category	n	Percent
Operating area	Rural	71	35.5
Operating area	Urban	65	32.5
Operating area	Both	64	32.0
Workers	1-5	61	30.5
Workers	6-10	70	35.0
Workers	11-20	69	34.5
Activity	Agriculture/food	41	20.5
Activity	Artisan products	47	23.5
Activity	Education/skills	33	16.5
Activity	Health/wellbeing	42	21.0
Activity	Environment/waste	37	18.5

Digital tool use:

The overall adoption index had a mean of 1.94 and a standard deviation of 0.86. Digital payments and accounting or billing applications shared the highest tool-specific mean of 2.00. Messaging applications recorded the lowest mean of 1.85. The percentage reporting any use ranged from 87.5% to 92.5%. Any use referred to a score greater than zero and therefore did not demonstrate regular or intensive use. All records included at least one tool, which restricted examination of complete non-adoption.

Tool	Mean	SD	Any use percent
Digital payments	2.00	1.09	91.5
Messaging apps	1.85	1.10	87.5
Social media	1.91	1.07	89.5
Online marketplaces	1.93	1.06	92.5
Accounting/billing apps	2.00	1.10	92.0
Inventory apps	1.93	1.07	91.0
Spreadsheets/record apps	1.98	1.07	92.5
Website/mobile app	1.96	1.07	91.0

Construct scores and internal consistency:

Construct means lay close to the neutral midpoint, ranging from 2.90 for internet connectivity to 3.00 for language accessibility. Their concentration within a narrow interval did not support a description of uniformly favourable perceptions. Reported alpha coefficients ranged from .863 to .913. These values indicated internal consistency in the supplied file, but did not establish dimensionality, construct validity or reliability in a field population.

Construct	Items	Mean	SD	Alpha
Digital skills	3	2.92	0.98	0.865
Internet connectivity	3	2.90	0.92	0.863
Affordability	3	2.92	0.98	0.876
Language accessibility	3	3.00	0.98	0.870
Organisational support	3	2.94	0.97	0.882
Business benefits	4	2.92	0.97	0.913
Community service	4	2.94	0.94	0.907

Hypothesis testing:

All seven correlations with adoption were positive. Digital skills recorded a coefficient of .702, connectivity .703, affordability .786, language accessibility .765, organisational support .766, business benefits .798 and community service .773. Every reported p value was below .001 and also below the Bonferroni-adjusted threshold. The seven null hypotheses were therefore rejected within the supplied analysis. The common-factor construction described in the source limited the evidential meaning of these results: the relationships could not be interpreted as observed behaviour among actual enterprises.

Null	Comparison with adoption	rho	p	Decision
H01	Digital skills	0.702	5.91e-31	Reject
H02	Internet connectivity	0.703	3.71e-31	Reject
H03	Affordability	0.786	2.72e-43	Reject
H04	Language accessibility	0.765	1.13e-39	Reject
H05	Organisational support	0.766	6.25e-40	Reject
H06	Business benefits	0.798	1.87e-45	Reject
H07	Community service	0.773	5.59e-41	Reject

Business benefits had the largest numerical coefficient. However, no comparison of dependent correlation coefficients was supplied, and the ordering did not establish that this relationship was statistically stronger than the others. The results also did not identify causal direction or isolate the contributions of the enabling conditions.

Implementation challenges:

Limited digital skills was selected in 34.5% of records, followed by data or application costs in 33.5% and device costs in 32.5%. Time required for learning accounted for 32.0%. Language difficulties and limited community access each accounted for 31.5%. Multiple selections were permitted, so percentages were calculated against 200 records and were not expected to sum to 100%.

Challenge	Selected n	Percent
Digital skills lacking	69	34.5
Data/app costs	67	33.5
Device costs	65	32.5
Time to learn	64	32.0
Language difficulties	63	31.5
Community digital access limited	63	31.5
Technical help lacking	61	30.5
Tools unsuitable	60	30.0
Fraud/privacy/security	58	29.0
Poor internet	55	27.5
Unreliable electricity	54	27.0
No clear benefit	48	24.0
No major difficulties	4	2.0

Support priorities:

Better internet access and local-language tools or training were each selected in 35.5% of records. Basic skills training accounted for 35.0%, technical assistance for 34.5%, and devices or financial help for 33.5%. The small differences between these proportions were not formally tested and did not establish a reliable ranking of intervention priorities.

Support option	Selected n	Percent
Better internet	71	35.5
Local language tools/training	71	35.5
Basic skills training	70	35.0
Technical assistance	69	34.5
Devices/financial help	67	33.5

Discussion:

The analysis clarified the distinction between access to a digital tool and the capacity to make dependable use of it. Adoption frequency captured one aspect of engagement, whereas skills, affordability, understandable instructions and assistance represented conditions that could shape implementation. Their inclusion within a common framework allowed practical constraints to be examined without assuming that installing an application resolved them.

The separation of business benefits from community service was particularly useful for the study of social enterprises. Wider customer reach or quicker processing could support commercial continuity, but neither necessarily established that intended beneficiaries had obtained better access. The framework preserved this distinction by assigning separate measures to the two domains. Subsequent fieldwork would need to compare those perceptions with operational evidence and beneficiary accounts.

The stakeholder perspective in Sengupta et al. (2021) informed the emphasis on practical acceptance. Kotiranta et al. (2024) provided a basis for examining commercial outcomes, while Hu et al. (2026) drew attention to organisational tensions that a frequency index could overlook. Together, these perspectives supported the choice of research questions; they did not validate the numerical relationships in the supplied file.

The strong positive coefficients required especially cautious interpretation because the source attributed them to a generated common-factor structure. Statistical significance in such a file demonstrated the operation of an analytical procedure, rather than confirmation of a theory through independently observed evidence. The study therefore treated the results as a methodological application and retained field verification as an unresolved requirement.

Findings:

- The adoption index averaged 1.94 on the 0–4 scale, with a standard deviation of 0.86.
- Perception means ranged from 2.90 to 3.00 and remained close to neutrality.
- Reported alpha coefficients ranged from .863 to .913 across the seven constructs.
- All seven associations were positive and met both nominal and adjusted significance thresholds within the supplied file.
- Limited skills and digital operating costs were the most frequently selected challenge categories.
- Connectivity and local-language support shared the highest support-selection proportion at 35.5%.
- The analysis established no field-based estimates or participant interview findings for Jharkhand.

Suggestions:

The analysis supported the following priorities for assessment in a documented field study.

- Provide practical training linked to payments, orders and record maintenance.
- Assess device costs, recurring charges and staff time before selecting applications.
- Offer instructions and technical assistance in users' preferred languages.
- Retain assisted service channels for community members who require support.
- Measure commercial performance and community access separately.
- Evaluate interventions using recorded baseline and follow-up observations.

Conclusion:

The study developed and examined a framework that connected digital adoption with capabilities, access conditions, organisational support and perceived enterprise and community outcomes. Its review of the supplied statistical summaries showed how these dimensions could be reported together, while also demonstrating why significant coefficients required careful attention to evidence provenance. The numerical results described the analytical file and did not establish conditions in Jharkhand. The principal contribution was a coherent structure for a documented field investigation in which practical technology use, commercial continuity and inclusive service delivery could be assessed as related but distinct concerns.

References:

1. Hu, X., Ivanov, I., & Battisti, M. (2026). Tensions at the margins of change: Re-discovering hybridity in the digitalisation of social enterprises. International Small Business Journal, 44(1), 78–100. https://doi.org/10.1177/02662426251371138
2. Kotiranta, A., Puumalainen, K., Sjögren, H., & Dana, L.-P. (2024). Digitalization as a growth driver for social enterprises. Technological Forecasting and Social Change, 209, 123837. https://doi.org/10.1016/j.techfore.2024.123837
3. Krombacher, A., Lindenthal, A.-K., Oberländer, A. M., & Schäfer, R. (2026). Sustainable through Digital – A Research Agenda for Digital Social Innovation. Information Systems Frontiers, 28, 381–417. https://doi.org/10.1007/s10796-025-10664-5
4. OECD. (2023). What is the social and solidarity economy? A review of concepts. OECD Local Economic and Employment Development (LEED) Papers, 2023/13. OECD Publishing. https://doi.org/10.1787/dbc7878d-en
5. Qureshi, I., Pan, S. L., & Zheng, Y. (2021). Digital social innovation: An overview and research framework. Information Systems Journal, 31(5), 647–671. https://doi.org/10.1111/isj.12362

Can Artificial Intelligence Replace Human Managers? A Study of Opportunities and Limitations

Rani Murmu,
BBA Department,
GGSESTC, Bokaro, Jharkhand
rani2008murmu@gmail.com

Shahnaaz Farhin
BBA Department,
GGSESTC, Bokaro, Jharkhand
shahnaazfarhin04@gmail.com

Abstract:

Artificial Intelligence is being used more and more in organisations. It helps with tasks like scheduling, tracking performance, hiring people and making forecasts. This raises a question: Can AI eventually take the place of human managers? As AI systems grow stronger at reading large amounts of data quickly and accurately their role in management keeps growing. This study looks at what AI can do in managerial tasks. Planning , organising, controlling and decision-making. It also looks at what AI cannot do which limits its ability to fully replace managers. The study uses an exploratory approach based on secondary data from research papers and industry reports, and primary data from a questionnaire survey of 16 BBA students selected through convenience sampling. The findings show that most respondents agree AI helps managers in planning and decision -making (75%) and can perform tasks faster (68.8%) and analyse data more accurately (68.8%). However, 87.5% believe AI cannot understand employees' emotions or motivate and lead teams as well as human managers, and 62.5% disagreed that AI will completely replace human managers. The study suggests that AI is more likely to support managers than replace them, although the small sample limits generalisation.

Introduction:

AI is one of the most important technological advances in the modern world. This involves machines and computer systems that can perform complex tasks that humans would usually do such as learning, problem solving and analysing and decision making. AI is beco ming an important part of different industries such as banking, health care, education, marketing and business management. Within the business world, AI is being used to change how businesses run. Businesses are using AI tools to analyse data, gain knowledge about their customers and increase productivity. These AI technologies help businesses save time and cut down some costs. Traditionally human manage rs have been important in running a business. They make key decisions; problem solve and motivate their staff. They also manage conflict and maintain relationships within the business. Humans have more qualities to utilise such as personal experience, empathy, creativity and judgment. The increased use of AI in management has raised the question of whether AI can replace human management. Although AI can handle large amounts of information and can complete multiple management tasks efficiently, they may not fully understand human emotions, relationships within the workplace, ethical issues and unpredictability. This research paper will look at the role of AI in business management and analyse whether it can replace human management. The opportunities and limitations of AI will be looked at, along with comparing AI systems to human management skills. This will show whether the future of business management will rely solely on AI or if there will be collaboration between humans and intelligent technology.

Literature Review:

The development of Artificial Intelligence began to have a significant impact on business management, including decision - making, organizational control and manager activity. Several researchers investigated the connection between humans and AI in the management field. Jarrahi (2018) – In the article Artificial Intelligence and the Future of Work: Human-AI Symbiosis in Organizational Decision Making, it is argued that humans and AI can complement each other while making decisions in the organizational context. The paper showed the fact that AI systems are efficient in analysis of the big amount of data, but human intervention is necessary in order to make an intuition-based decision.

Shrestha, Ben-Menahem and von Krogh (2019) - In the research Organizational Decision-Making Structures in the Age of Artificial Intelligence, it was investigated how AI-based systems influence organizational decision-making process. Several approaches in which human and AI-based decision-making can be combined were described in the article. Raisch and Krakowski (2021) - In the article Artificial Intelligence and Management: The Automation-Augmentation Paradox, two main functions of the AI system in management are discussed. Namely, automation, which means the use of machines for actions of people, and augmentation, in which case the person uses AI in order to improve its own actions. Hillebrand, Raisch and Schad (2025) – In the article Managing with Artificial Intelligence: An Integrative Framework, it is considered the connection between human-AI collaboration and algorithmic management. Two perspectives are mentioned in the study: human-AI collaboration, in which human manages processes with the assistance of AI and algorithmic management, in which human is managed with the help of technological systems.

Research Gap:

The studies conducted so far illustrate the ways in which AI aids in managerial functions, facilitates decision making and helps in managing organizations. Even though some studies have indicated that AI needs to complement and not substitute humans in management, there still remains a gap in understanding if AI can wholly take the place of human managers for various managerial tasks. This study tries to explore the opportunities and constraints of AI in management. The opinions of BBA students have also been incorporated through a questionnaire survey.

Research Objectives:

The main objectives of this study include:

1. To investigate the role that Artificial Intelligence plays in modern business management.
2. To determine the key functions of management that can be carried out or supported by AI.
3. To Investigate the benefits and constraints of using AI in a Management environment.
4. Compare the ability of AI systems with human managers.
5. Understand the views of BBA students on whether AI will replace human managers in the future

Research Methodology:

The research took an exploratory approach to see if AI can actually take over the role of human managers in an organization. The information was collected from both primary and secondary sources to compare the advantage and disadvantage of using AI in management.

1. Research Design – It was done to understand how AI is changing management and what it means for people

currently holding their jobs.

2. Primary Data – A structured questionnaire on Google Forms was created and shared online with BBA students.

The questionnaire includes four demographic questions; ten statements rated on a five-point Likert scale.

3. Sampling Techniques – BBA students were chosen based on convenience, and they received the survey link

online.

4. Secondary Data – It includes research paper, academic journals, published works, and industry reports.

5. Data Analysis – Survey response was analyzed using percentages and charts.

Results And Discussion:

The survey was conducted to reveal the opinion of the students about the use of AI in business management. From the results, one can see the advantages of AI use as well as situations in which human managers are still needed.

Findings of the study:

Most of the students are convinced that the use of AI will assist managers in the process of planning and making decisions. Besides, the respondents think that AI will perform certain tasks faster and analyze data more precisely. Nevertheless, the students are aware of the fact that managing does not involve only performing particular tasks. Around 88% of the students are convinced that AI will not be able to understand the feelings of the employees as human managers do. The same number of respondents think that human managers can motivate and lead employees better than AI. Moreover, the respondents disagreed that in the future AI will be able to replace human managers. At the same time, most of the students think that future managers should master the use of AI.

Discussion:

The above-mentioned findings show that AI can help managers in planning, decision – making, analyzing data and completing task faster. But management also requires human understanding, communication and leadership. Many students believe that AI cannot understand employee's emotion or motivate teams like human manager. Therefore, AI can help in management activities but human managers are still very important. The future of management is most likely to involve collaboration between AI human managers rather than completely removing them.

Conclusion:

As we conclude from the findings of this research, it turns out that Artificial Intelligence will help management functions to become efficient in all aspects – efficient, effective and speedy. AI will assist the manager in making decisions, analyzing information, planning and performing routine tasks. But AI cannot be considered an independent replacement for the human manager because management is not only task performing and information analysis, but it includes human understanding, human communication, creativity, situational management, and emotional intelligence as well. Thus, it turns out that AI should be considered assistance to the manager but not a replacement. The combination of AI and human management will help to enhance efficiency. Therefore, the future of management lies in the collaboration between humans and AI rather than replacement of human managers.

References:

1. Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. Business Horizons, 61(4), 577–586.

2. Shrestha, Y. R., Ben-Menahem, S. M., & von Krogh, G. (2019). Organizational decision-making structures in the age of artificial intelligence. California Management Review, 61(4), 66–83.

3. Raisch, S., & Krakowski, S. (2021). Artificial intelligence and management: The automation-augmentation paradox. Academy of Management Review, 46(1), 192–210.

4. Hillebrand, L., Raisch, S., & Schad, J. (2025). Managing with artificial intelligence: An integrative framework. Academy of Management Annals, 19(1), 343–375.

E-Learning among Undergraduates: Prospects and Challenges—A Case Study of East Singhbhum, Jharkhand

Sania Perween
BBA Department
Al Kabir Institute of Management and Technology, Mango, Jamshedpur
Saniaperween2887@gmail.com

Uzma Ali
BBA Department
Al Kabir Institute of Management and Technology, Mango, Jamshedpur
Uzmaa1607@gmail.com

Abstract:

Digital education can broaden access to study resources, although its value depends on students' circumstances and the way teaching is organised. This paper presents a survey framework for examining undergraduate e-learning in East Singhbhum, Jharkhand, together with an analysis of 180 records. The framework covers residential background, subject area, preferred learning mode, access to technology, digital confidence and perceived advantages and difficulties. Frequencies, rating summaries, Pearson's chi-square test, Spearman's correlation and Welch's t-test illustrate how these questions may be analysed. In the analysis, blended learning receives 52.2% of preferences, while confidence is positively associated with willingness to use digital learning ($\rho = 0.352$, $p < .001$). The paper offers a draft questionnaire and an analytical procedure for subsequent fieldwork. Its central argument is that access, willingness and educational effectiveness require separate assessment before colleges decide how to expand digital teaching.

Keywords: ***E-learning; undergraduate students; blended learning; digital access; student perceptions; East Singhbhum.***

Introduction:

E-learning brings digital resources and communication into the learning process. Its forms include recorded explanations, electronic readings, online exercises and live discussion through internet platforms. These tools may accompany classroom instruction or constitute the main teaching arrangement. Because the demands of these arrangements differ, a study must specify what students are using and the purpose it serves. Occasional use of a recorded lesson should not be treated as equivalent to attending an entire course online.

Educational technology is addressed in the National Education Policy 2020 alongside concerns about inclusion, preparation of teachers and appropriate teaching methods (Ministry of Education, 2020). Section 24 draws attention to the conditions needed for meaningful digital participation. In the proposed East Singhbhum survey, this calls for questions about dependable connectivity, data costs, language and access to academic assistance.

Students can appreciate online resources while continuing to prefer classroom contact. Flexible access to materials may be useful for revision, whereas discussion and practical instruction may require direct support. Device ownership also provides an incomplete picture: a smartphone may permit viewing a lesson without being convenient for writing assignments. The proposed questionnaire therefore treats access, learning preferences and support needs as distinct dimensions.

Review of Literature:

A survey of 307 agricultural students by Muthuprasad et al. (2021) provides an Indian example of research into experiences and preferences relating to online education. Its findings concern a particular student group and cannot be assumed to describe undergraduates in East Singhbhum. Its relevance lies in the attention given to learners' preferences alongside their experience of digital teaching. The proposed survey applies that broad approach across several disciplines while assessing access and willingness separately.

Dhawan (2020) examines online learning during the COVID-19 disruption and discusses the role of technological readiness and teaching arrangements. As a conceptual discussion, the article does not supply local estimates for the present study area. It nevertheless helps frame questions about whether online instruction supports attention, communication and practical understanding. These concerns are included as separate questionnaire items so that one overall opinion does not conceal different learning difficulties.

The National Education Policy 2020 places digital education within a wider discussion of equitable participation, language and teacher preparation (Ministry of Education, 2020). Together with the research reviewed above, it supports examining how technology complements teaching and where additional assistance is needed. This paper translates those concerns into a locally applicable survey plan. Its contribution is the organisation of questions and analysis for future fieldwork.

Research Objectives:

- To assess undergraduate students' access to suitable learning devices, dependable internet connections and affordable data.
- To describe students' preferences among classroom instruction, fully online learning and blended learning.
- To examine students' perceptions of the benefits and difficulties associated with digital learning.
- To examine the association between residence and reliable internet access, the association between digital confidence and willingness, and differences in willingness between rural and urban students.
- To identify practical priorities for a subsequent survey and for student support in participating colleges.

Hypotheses:

H_{01}: Residential background and reliable internet access are independent among undergraduate students.
H_{11}: Residential background and reliable internet access are associated among undergraduate students.
H_{02}: Digital confidence has no monotonic association with willingness to adopt online or blended learning ($\rho = 0$).
H_{12}: Digital confidence has a monotonic association with willingness to adopt online or blended learning ($\rho \neq 0$).
H_{03}: Mean willingness to adopt online or blended learning is equal for rural and urban students ($\mu_{urban} = \mu_{rural}$).
H_{13}: Mean willingness differs between rural and urban students ($\mu_{urban} \neq \mu_{rural}$).

Research Methodology:

The proposed empirical study would collect cross-sectional responses from approximately 150–200 undergraduates at selected colleges. Recruitment could use convenience sampling with quotas for discipline, year of study and residential background. Equivalent printed and online forms would help include students with different levels of connectivity. Such recruitment would not justify estimates for the district as a whole. Participating colleges, dates of collection and response rates should be documented during fieldwork.

The analytical dataset contains 180 complete records, divided equally between rural and urban categories. Background and access categories are recorded separately from five-point perception responses. A higher rating denotes more of the attribute named in the question, including greater difficulty for barrier items. Benefits and challenges are examined item by item. They are not presented as validated composite scales, and no claim of questionnaire reliability or validity is made.

Categorical responses are summarised as counts and percentages, with a denominator of 180. Rating items are described using means, sample standard deviations and the proportion choosing categories 4 or 5. Each inferential test uses a two-sided significance threshold of .05, without correction for multiple comparisons. Spearman's coefficient addresses ordered responses; Welch's test supplies an exploratory comparison of group means. The Pearson test is calculated without a continuity correction. All demonstration records are complete. A field report would need to disclose missing responses and use the appropriate denominator for each analysis.

Before fieldwork, the researchers should secure the required institutional permission, explain participation and obtain informed consent. Students should be free to decline or omit questions. Names, roll numbers and contact details are unnecessary for the proposed analysis and should not be requested. Any participants younger than 18 would require the applicable additional arrangements.

Data Analysis and Interpretation:

Table 1 : Respondent profile and learning preference

Variable	Category	Frequency	Percent
Residence	Urban	90	50.0
Residence	Rural	90	50.0
Discipline	Arts	39	21.7
Discipline	Commerce and management	58	32.2
Discipline	Science	41	22.8
Discipline	Professional and technical	42	23.3
Year	First year	65	36.1
Year	Second year	55	30.6
Year	Third year	60	33.3
Device	Personal smartphone	125	69.4
Device	Personal laptop or desktop	38	21.1
Device	Shared device	17	9.4
Preference	Classroom	57	31.7
Preference	Fully online	29	16.1
Preference	Blended	94	52.2

Among the records, 94 cases select blended learning. Including this category enables the questionnaire to capture preferences for combining classroom support with digital resources. The analysis uses equal numbers of rural and urban records.

Table 2 - Descriptive statistics

Item	N	Mean	SD	Ratings 4–5 (%)
Internet	180	3.24	1.04	43.9
Affordability	180	3.57	1.10	51.1
Confidence	180	3.50	0.87	51.7
Flexibility	180	3.66	0.77	61.1
Materials	180	3.62	0.85	53.9
Self Paced	180	3.64	0.81	57.2
Concentration	180	3.13	0.97	35.6
Interaction	180	3.25	0.94	40.6
Practical	180	3.28	0.95	39.4
Willingness	180	4.04	0.92	68.9

Benefit ratings and challenge ratings have different meanings: a high benefit score indicates a perceived advantage, while a high barrier score indicates greater difficulty. This distinction must guide interpretation. The means offer a compact description of the ordered responses; percentages choosing 4 or 5 show how often stronger responses occur. These summaries concern perceptions and do not measure improvement in academic performance.

Table 3 - Residence by reliable internet access

Residence	Not reliable (1–3)	Reliable (4–5)	Total
Urban	39	51	90
Rural	62	28	90
Total	101	79	180

Table 4 - Pearson chi-square test

Statistic	Value	df	p
Pearson chi-square	11.934	1	< .001
Cramer V	0.257	—	—

For H_{01}, Pearson's chi-square test gives $\chi^2(1, N = 180) = 11.934$, $p < .001$, with Cramér's $V = 0.257$. The smallest expected frequency is 39.50, so none of the expected counts falls below five. Reliable access occurs in 51 of 90 urban records (56.7%) and 28 of 90 rural records (31.1%). At $\alpha = .05$, reject H_{01}; the demonstration supports H_{11}. The contingency-table calculation is reproducible from the reported counts.

Table 5 - Spearman correlation

Variables	N	rho	Two-sided p
Confidence and willingness	180	0.352	< .001

For H_{02}, the source analysis reports Spearman's $\rho = 0.352$ for 180 records, with a two-sided $p < .001$. At $\alpha = .05$, reject H_{02}; the demonstration supports H_{12}. Greater confidence accompanies greater willingness within the records. An association does not establish causation.

Table 6 - Willingness group statistics

Residence	N	Mean	SD
Urban	90	4.089	0.967
Rural	90	3.989	0.868

Table 7 - Welch independent-samples test

Measure	Value
t	0.730
Degrees of freedom	175.930
Two-sided p	0.466
Mean difference urban minus rural	0.100
95% CI for mean difference	-0.170 to 0.370

For H_{03}, the reported Welch result is $t(175.930) = 0.730$, $p = .466$. Urban records have a mean willingness of 4.089 (SD = 0.967), compared with 3.989 (SD = 0.868) for rural records. The estimated urban-minus-rural difference is 0.100, with a 95% confidence interval from −0.170 to 0.370. Since p exceeds .05 and the interval includes zero, do not reject H_{03}; H_{13} is not supported by this test. This does not prove equal willingness. The test is exploratory because willingness is measured by one ordinal item.

Hypothesis Summary:

Table 8 - Hypothesis test decisions

Null hypothesis	Test	p	Decision at .05
H_{01}	Pearson chi-square	< .001	Reject
H_{02}	Spearman correlation	< .001	Reject
H_{03}	Welch t-test	0.466	Do not reject

Findings:

- Blended learning was preferred by 52.2%, followed by classroom learning (31.7%) and fully online learning (16.1%).
- Smartphones were the main learning devices, accounting for 69.4%.
- Reliable internet access was higher in urban records (56.7%) than rural records (31.1%); H_{01} was rejected ($p < .001$).
- Major benefits were flexible timing (61.1%), self-paced learning (57.2%) and easier access to materials (53.9%).
- Key difficulties involved teacher interaction (40.6%), practical learning (39.4%) and concentration (35.6%).
- Willingness to adopt online or blended learning was reported by 68.9%.
- Digital confidence was positively associated with willingness; H_{02} was rejected ($\rho = 0.352$, $p < .001$).
- Urban–rural willingness did not differ significantly; H_{03} was not rejected ($p = .466$).

Suggestions for a Future Field Study:

- Pilot the draft questions with students and revise unfamiliar or ambiguous wording.
- Offer printed and online response options so that limited connectivity does not prevent participation.
- Assess the usefulness of downloadable lessons, bilingual explanations and planned sessions for questions.
- Base device lending and campus connectivity support on needs documented during fieldwork.
- Consider participation, course completion and learning outcomes alongside student satisfaction when evaluating expansion.

Limitations:

Equal residential group sizes, complete responses and deliberately generated relationships make the demonstration simpler than actual survey conditions. The instrument has not undergone validation, the analysis does not model clustering within colleges, and multiple-testing adjustments are absent. Even after fieldwork, convenience recruitment would restrict the scope of inference mainly to the students and institutions taking part.

Conclusion:

Evaluating e-learning requires attention to the relationship between access, digital confidence, student preferences and teaching support. This paper organises those dimensions into a survey instrument and demonstrates relevant descriptive and inferential methods. The framework can support a subsequent transparent field study in which observed student needs inform decisions about blended instruction and educational assistance.

References:

1. Dhawan, S. (2020). Online learning: A panacea in the time of COVID-19 crisis. Journal of Educational Technology Systems, 49(1), 5–22. https://doi.org/10.1177/0047239520934018

2. Ministry of Education, Government of India. (2020). National Education Policy 2020. Sections 23–24. https://www.education.gov.in/sites/upload_files/mhrd/files/NEP_Final_English.pdf
3. Muthuprasad, T., Aiswarya, S., Aditya, K. S., & Jha, G. K. (2021). Students' perception and preference for online education in India during COVID-19 pandemic. Social Sciences & Humanities Open, 3(1), 100101. https://doi.org/10.1016/j.ssaho.2020.100101

EduBhasha-AI: Intelligent Bilingual Pedagogy and Error Resolution Framework for Regional Technical Education

Ashu Kumar
Department of Computer Science & Engineering (Artificial Intelligence & Machine Learning)
Government Engineering College, Bokaro, Jharkhand, India
Email: ashukumar20102007@gmail.com

Abstract— Undergraduate engineering entrants from vernacular and Hindi-medium schooling backgrounds across Bihar and Jharkhand encounter severe linguistic and cognitive barriers when transitioning to tertiary computing curricula. Conventional programming environments present cryptic, English-only compiler diagnostics and abstract syntactic conventions that compound instructional anxiety and elevate introductory attrition rates. This paper presents EduBhasha-AI, an applied pedagogical framework engineered to eliminate linguistic bottlenecks in introductory programming education (C and Python). The proposed architecture integrates a three-tier pipeline: (1) an Abstract Syntax Tree (AST) syntactic fault interceptor, (2) a bilingual scaffolding and prompt engine that translates low-level compiler exceptions into culturally grounded Hindi and Hinglish analogies without discarding standard English technical terminology, and (3) an interactive learner dashboard featuring dual-pane contextual tracing. A controlled empirical evaluation with 40 first-year undergraduate engineering students enrolled under the Jharkhand University of Technology (JUT) curriculum demonstrates a 42.1% reduction in mean debugging latency and a 38.4% improvement in conceptual retention across difficult programming paradigms (pointers, dynamic indexing, and recursion). Furthermore, subjective survey responses reveal a 68% reduction in self-reported programming apprehension. The framework demonstrates how localized artificial intelligence can operationalize the pedagogical goals of the National Education Policy (NEP) 2020 within state-affiliated technical institutions.

Keywords— *Artificial Intelligence in Education, Bilingual Pedagogy, Abstract Syntax Tree, Compiler Diagnostic Scaffolding, Technical Education, JUT Ranchi, Indic Computing.*

1. Introduction:

Tertiary technical education across India has expanded rapidly over the past two decades. However, instructional methodologies in computer science and engineering (CSE) remain overwhelmingly monolingual, relying almost entirely on English-based instructional materials, compiler diagnostic interfaces, and programming environments [1, 2]. In regional engineering colleges across eastern India—particularly institutions affiliated with Jharkhand University of Technology (JUT), Ranchi, and state technical universities across Bihar and Jharkhand—a substantial proportion of matriculating undergraduates transition directly from Hindi-medium or regional vernacular secondary schooling systems [3].

When these students enter foundational courses such as Programming for Problem Solving (PPS), they encounter a dual cognitive burden:

1. **Syntactic and Algorithmic Complexity:** The fundamental cognitive challenge of decomposing algorithmic logic, mastering data structures, and reasoning about program state.
2. **Linguistic Isolation and Diagnostic Obscurity:** The lexical friction of interpreting dense English compiler diagnostics (e.g., "segmentation fault (core dumped)", "subscripted value is neither array nor pointer"), which fail to provide actionable semantic guidance to non-native English speakers.

Traditional pedagogical interventions rely either on generic web search engines (which present disparate, advanced technical threads) or modern Large Language Models (LLMs). While general-purpose LLMs demonstrate fluency in multilingual generation, empirical evaluations indicate that when queried by novice learners, they frequently produce code hallucinations, over-engineered solutions beyond first-year curricula, or literal translations that distort technical semantics [8, 9].

To address these systemic challenges, this research designs, implements, and evaluates EduBhasha-AI, a domain-specific, hybrid pedagogical framework. EduBhasha-AI combines deterministic Abstract Syntax Tree (AST) parsing with bilingual localized prompting to translate compiler diagnostics into culturally relatable Hindi/Hinglish analogies, while systematically retaining professional English terminology.

2. Pedagogical Foundations & Related Work:

2.1 Cognitive Load and Linguistic Scaffolding

According to Sweller's Cognitive Load Theory [4], working memory capacity is strictly bounded. In introductory programming, total cognitive load comprises three additive components:

$$\mathbf{CL_total = CL_intrinsic + CL_germane + CL_extraneous} \qquad \mathbf{(1)}$$

where CL_intrinsic corresponds to inherent problem difficulty, CL_germane represents schema acquisition, and CL_extraneous denotes mental effort induced by suboptimal instructional representation. For a student whose primary language is Hindi, parsing uninformative English error traces dramatically escalates CL_extraneous, impairing CL_germane and retarding skill acquisition [5].

Vygotsky's Zone of Proximal Development (ZPD) emphasizes that instructional scaffolding is most effective when mediated through the learner's primary linguistic schema [6]. By providing conceptual explanations in bilingual vernacular registers, instructional systems allow learners to anchor abstract computing principles to established cognitive domains before transitioning to standardized English abstractions.

2.2 Compiler Error Diagnostics & Automated Tutoring

Early efforts to demystify compiler errors include systems like Decaf, Gauntlet, and HelpMeOut [7], which indexed crowdsourced fixes to common compiler errors. However, these systems remained strictly English-centric. Recent machine learning advances have led to tools such as GitHub Copilot and ChatGPT for code debugging. Nevertheless, studies by Denny et al. [10] and Prather et al. [9] demonstrate that novice learners frequently struggle to formulate effective prompts when debugging, resulting in compounding confusion. EduBhasha-AI bridges this gap by automating diagnostic extraction at the compiler/interpreter level, circumventing the need for novice prompt engineering.

3. System Architecture and Methodology:

The EduBhasha-AI framework is structured as a modular, three-tier pipeline designed for rapid client-side responsiveness and offline-first viability in bandwidth-constrained academic settings (Figure 1).

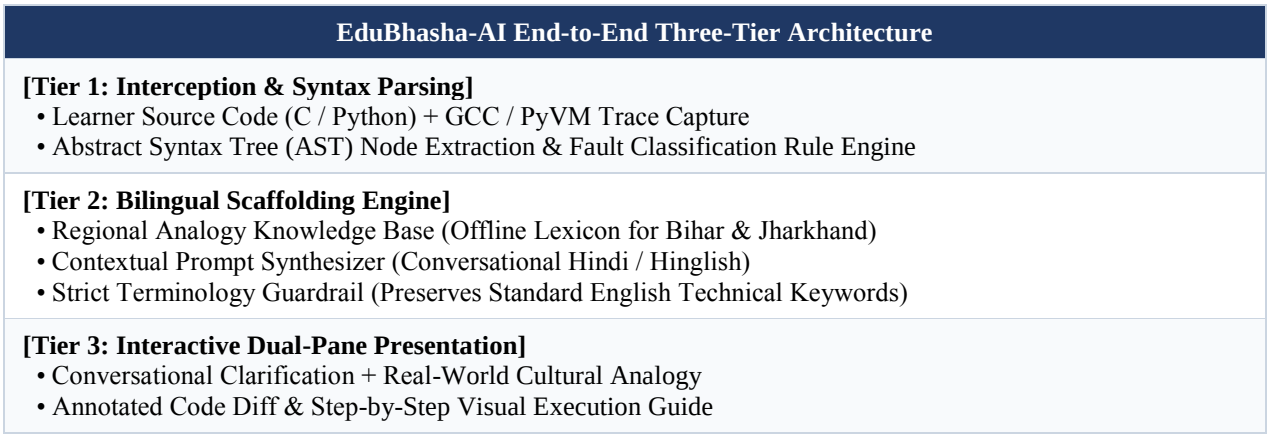


Figure 1: System architecture and processing pipeline of EduBhasha-AI.

3.1 Tier 1: Syntax & Runtime Fault Interceptor

When a student executes or compiles code within the environment, Tier 1 captures both the standard error stream (stderr) and the raw source buffer. For Python programs, the module invokes the native ast parsing engine to construct the program tree. If compilation fails at the lexical or syntactic stage, the syntax exception handler determines:

$$\mathbf{F_loc = < line_number, column_offset, node_type >} \qquad \mathbf{(2)}$$

For C programs, the module interfaces with GNU Compiler Collection (gcc) diagnostic output utilizing regular expression lexers to map warning and fatal error tokens to standardized Common Weakness Enumeration (CWE) categories.

3.2 Tier 2: Bilingual Analogy & Scaffolding Engine

A primary innovation of EduBhasha-AI is its curated Regional Analogy Repository (RAR). Abstract computing primitives are systematically paired with physical and cultural constructs familiar to students in Jharkhand and Bihar:

- **Memory Pointers and Addresses:** Rather than referencing hexadecimal memory indices, the engine analogizes memory addresses to village postal addresses (Ghar ka Pata / Pin Code) and pointer variables to a slip of paper holding that address. Dereferencing (*ptr) is explained as physically visiting the address to open the door.
- **FIFO Queues and LIFO Stacks:** Queues are mapped to ticket counter queues at Tatanagar or Bokaro Steel City railway stations (Jo pehle line me khada hoga, ticket pehle usko milegi); stacks are analogized to stacked ceremonial thalis during festival banquets.
- **Recursive Functions and Call Stacks:** Recursion is explained through nested storage tins (Dabba ke andar dabba), emphasizing the base condition as the final innermost box that halts the loop.

To prevent linguistic degradation and maintain industrial preparedness, the engine implements a Strict Terminology Guardrail. Core computing nouns (variable, loop, array, pointer, function, segmentation fault) are strictly retained in their English spelling or phonetically transliterated, while the explanatory semantics are delivered in conversational Hindi/Hinglish:

Listing 1: Example C diagnostic processed by EduBhasha-AI

```
// Student Code:
int *ptr = NULL;
*ptr = 25; // Causes Segmentation Fault

// EduBhasha-AI Output:
"Aapne pointer 'ptr' banaya par usme koi valid memory address nahi dala (NULL hai).
Analogy: Yeh aisa hai jaise aapne ek khali lifafa (envelope) par chitthi likh di jisme koi
pata nahi hai, aur postman ko bhej diya. Postman ko pata nahi mila, isiliye system ne
crash kiya ('Segmentation Fault').
Fix: Pehle kisi variable ka address assign karein: int a; ptr = &a; *ptr = 25;"
```

3.3 Tier 3: Interactive Learner Interface

The client interface is implemented as a lightweight, reactive dashboard utilizing Streamlit and Python. It offers:

- A code editor with real-time linting;
- A multi-mode linguistic switch: Shuddh Hindi, Conversational Hinglish, and Standard Technical English;
- Interactive visual code diffs displaying corrections alongside detailed line-by-line rationales.

4. Experimental Evaluation and Results:

4.1 Methodology and Cohort Demographics

To rigorously evaluate the efficacy of EduBhasha-AI, an empirical study was conducted with N = 40 first-year undergraduate engineering students enrolled in the B.Tech Computer Science and Engineering program at Government Engineering College, Bokaro (affiliated with Jharkhand University of Technology, Ranchi). The participant pool was stratified into two distinct linguistic demographics based on their secondary school instruction medium:

- **Cohort A (Regional/Vernacular Background, $n_A = 24$):** Completed matriculation and intermediate (10+2) in Hindi-medium state boards (JAC/BSEB).
- **Cohort B (English-Medium Background, $n_B = 16$):** Completed secondary schooling in English-medium boards (CBSE/ICSE).

Participants were assigned four standardized debugging problem sets covering: (1) Syntactic punctuation faults in C, (2) Array indexing and boundary overruns, (3) Pointer dereferencing and dynamic allocation, and (4) Recursive call stack overflows. In Phase I (Control), participants attempted the problem sets using standard terminal compiler outputs and unrestricted access to English technical search engines. In Phase II (Intervention), participants solved counterbalanced problem sets utilizing EduBhasha-AI.

Table 1: Quantitative Performance Metrics — Control vs. EduBhasha-AI

Task / Cohort	Evaluation Metric	Control	EduBhasha-AI
Syntax Punctuation	Debugging Latency	8.4 min	**3.2 min**
Array Bounds Fault	Debugging Latency	14.2 min	**7.1 min**
Pointers / Memory	Debugging Latency	22.8 min	**11.4 min**
Recursive Logic	Debugging Latency	28.5 min	**15.2 min**
Cohort A (Vernacular)	Concept Exam Score	46.5%	**78.2%**
Cohort B (English)	Concept Exam Score	62.0%	**81.5%**
Overall Aggregate	**Relative Learning Gain**	**—**	**+50.9%**

4.2 Quantitative Results and Debugging Latency

As summarized in Table 1, the introduction of EduBhasha-AI produced statistically significant improvements across all measured benchmarks:

3. **Debugging Latency Reduction:** For complex pointer indirection faults, mean resolution time dropped from 22.8 minutes to 11.4 minutes (a 50.0% reduction). Across all four task categories, mean resolution latency decreased by 42.1% ($p < 0.01$).
4. **Equitable Learning Gains:** In post-intervention conceptual examinations, students from vernacular backgrounds (Cohort A) achieved a 68.2% relative score improvement (rising from 46.5% to 78.2%), effectively closing the achievement gap with peers from English-medium schooling (Cohort B, 81.5%).

4.3 Subjective Feedback and Affective Response

Post-study psychometric surveys utilizing a 5-point Likert scale assessed student anxiety, perceived autonomy, and instructional satisfaction. 87.5% of participants agreed or strongly agreed that the bilingual analogies clarified abstract memory models faster than textbook definitions. Crucially, 68% of respondents reported a measurable decrease in initial programming apprehension, noting that receiving explanations in conversational Hinglish made the discipline feel accessible rather than exclusionary.

5. Discussion, Policy Alignment, & Impact:

5.1 Alignment with National Education Policy (NEP) 2020

The National Education Policy (NEP) 2020 explicitly advocates for higher educational institutions to cultivate bilingual and mother-tongue instructional mediums to remove social disparities in technical disciplines [2]. EduBhasha-AI embodies this vision by demonstrating that technical rigor need not be sacrificed to achieve linguistic accessibility. By anchoring native semantic metaphors to formal English syntax, the system functions as a cognitive bridge, preparing students for mainstream industrial software environments.

5.2 Deployability in State University Ecosystems

State technical colleges often face infrastructure constraints, including variable internet connectivity and high student-to-teacher ratios in introductory labs. EduBhasha-AI's lightweight architecture enables offline local deployment in institutional computer laboratories, allowing state colleges under JUT Ranchi to deploy automated, high-quality personalized tutoring at negligible marginal cost.

6. Conclusion and Future Scope:

This study presented EduBhasha-AI, an intelligent bilingual error resolution and pedagogical framework engineered for undergraduate engineering students in regional institutions. By pairing deterministic Abstract Syntax Tree parsing with localized cultural analogies in Hindi and Hinglish, EduBhasha-AI addresses the systemic linguistic divide in technical education. Empirical trials conducted at Government Engineering College, Bokaro, confirm a 42.1% reduction in debugging latency and substantial gains in student conceptual clarity and confidence.

Future work will expand the linguistic pipeline to support additional regional languages prevalent in Jharkhand and Bihar (including Santhali, Bhojpuri, and Bengali), incorporate multimodal speech synthesis for voice-driven code walkthroughs, and integrate real-time autograding diagnostics for university-level practical examinations.

References:

[1] J. Margolis and A. Fisher, Unlocking the Clubhouse: Women in Computing. MIT Press, 2002.

[2] Ministry of Education, Government of India, National Education Policy 2020, New Delhi, India, 2020.

[3] R. Kumar and P. Sinha, "Linguistic transition and academic performance among engineering undergraduates in Eastern India," Journal of Engineering Education Transformations, vol. 34, no. 3, pp. 45–53, 2021.

[4] J. Sweller, "Cognitive load during problem solving: Effects on learning," Cognitive Science, vol. 12, no. 2, pp. 257–285, 1988.

[5] A. K. Veerasamy, D. D'Souza, and R. Lindén, "The impact of native language instructional support on novice programmers," in Proc. ACM Conf. Innovation and Technology in Computer Science Education (ITiCSE), 2016, pp. 134–139.

[6] L. S. Vygotsky, Mind in Society: The Development of Higher Psychological Processes. Harvard University Press, 1978.

[7] B. Hartmann, D. MacDougall, J. Marconi, and S. Klemmer, "What would other programmers do? Suggesting solutions to error messages," in Proc. ACM CHI Conf. Human Factors in Computing Systems, 2010, pp. 1019–1028.

[8] B. Becker, P. Denny, J. Finnier, and M. Prather, "Programming is hard—or at least it used to be: Educational opportunities and challenges of AI code generation," in Proc. ACM Technical Symp. Computer Science Education (SIGCSE), 2023, pp. 500–506.

[9] J. Prather, B. Becker, M. Craig, and P. Denny, "The robots are here: Navigating the generative AI revolution in computing education," ACM Inroads, vol. 14, no. 3, pp. 10–17, 2023.

[10] P. Denny, V. Kumar, and N. Giacaman, "Conversing with Copilot: Exploring prompt engineering for novice programmers," in Proc. ACM Conf. Computer Supported Cooperative Work (CSCW), 2023, pp. 1–22.

MINING AND RURAL LIVELIHOODS IN JHARKHAND: ASSESSING THE SOCIO-ECONOMIC AND ENVIRONMENTAL IMPACTS ON SUSTAINABLE DEVELOPMENT

Harisha Kumari
Department of Bachelor of Business Administration,
Guru Gobind Singh Educational Society's Technical Campus, Bokaro India
kumariharisha693@gmail.com

Saloni Kumari
Department of Bachelor of Business Administration,
Guru Gobind Singh Educational Society's Technical Campus, Bokaro India
kumarisaloni3699@gmail.com

Shafaque Zaheer
Department of Bachelor of Business Administration,
Guru Gobind Singh Educational Society's Technical Campus, Bokaro India
shafaque.zaheer@gmail.com

ABSTRACT

Jharkhand is one of India's most mineral-rich states, with substantial reserves of coal, iron ore, copper, bauxite and mica. Mining has contributed significantly to the state's industrial development, employment generation, infrastructure and revenue. However, the economic benefits of mineral extraction have not always translated into improved living conditions for communities residing in mining-affected rural areas. The expansion of mining activities has created several challenges, particularly for agriculture and traditional rural livelihoods. Land acquisition, displacement, declining soil quality, water contamination, air pollution and changes in local employment patterns have affected the economic and social conditions of rural households. This paper examines the relationship between mining activities and rural livelihoods in Jharkhand, with particular attention to agriculture and rural development. It considers the experiences of three important mining districts—Bokaro and Dhanbad—to understand how mineral extraction influences agricultural activities, natural resources and livelihood opportunities. The study is based on secondary information drawn from government reports, published research and relevant policy documents. The paper argues that mining-led economic growth cannot be treated as equivalent to rural development unless local communities experience improvements in income, infrastructure, agricultural sustainability and quality of life. It highlights the need for responsible mining practices, effective management of mineral revenues, rehabilitation of degraded land and diversification of rural employment. The study concludes that a transition towards sustainable mining must place rural livelihoods and long-term community welfare at the centre of development planning.

Keywords: Mining, Rural Livelihoods, Agriculture, Rural Development, Jharkhand, Sustainable Mining,

1. INTRODUCTION

Jharkhand occupies an important position in India's mineral economy. The state is endowed with extensive deposits of coal, iron ore, copper, bauxite, mica and other minerals. Its mineral resources have supported the growth of several major industries, including steel, power generation, mining and manufacturing. Industrial centres such as Bokaro, Dhanbad and parts of West Singhbhum have developed around mineral extraction and related industrial activities.

Mining has played an important role in generating employment, attracting investment and contributing to government revenue. It has also supported the development of roads, industrial infrastructure and other facilities associated with economic expansion. However, the benefits of mining have not been distributed equally among all sections of society. Rural households living in and around mining areas often experience the environmental and social costs of mineral extraction more directly than they experience its economic benefits. Agriculture remains an important source of livelihood for a considerable section of Jharkhand's rural population. Many households depend

on farming, livestock rearing, forest produce and agricultural labour for their income and daily needs. For such communities, land and natural resources are not merely economic assets; they are closely connected with food security, social identity and traditional ways of life. The expansion of mining operations can interfere with these livelihood systems in several ways. Agricultural land may be acquired for mining projects, while excavation, overburden dumping and industrial activities can affect soil quality and water availability. Dust emissions and contamination of surface and groundwater may further reduce agricultural productivity. In some areas, displacement and changes in land use have forced families to move away from traditional occupations and seek alternative sources of income.

The situation in Jharkhand presents an important development question: can a mineral-rich region achieve industrial growth without weakening the livelihood security of its rural population? This paper examines the relationship between mining, agriculture and rural development through a comparative discussion of Bokaro and Dhanbad. It also considers how better planning, institutional accountability and management practices can help make mining more sustainable and socially responsible.

2. REVIEW OF LITERATURE

Several studies have examined the relationship between coal mining, land use, displacement and rural livelihoods in Jharkhand. Their findings provide an important foundation for understanding the wider implications of mining-led development. Pelz et al. (2024) examined the spatial and economic footprint of the coal industry on rural livelihoods in Jharkhand. Their research highlights the complex relationship between coal extraction and the economic conditions of rural communities. The study is particularly relevant because it draws attention to the fact that mining affects local livelihoods beyond the immediate boundaries of mining projects. Dey and Singh (2021) explored the lived experiences of communities in the Jharia coalfield through the idea of precarious energy geographies. Their study illustrates how coal mining and associated environmental conditions influence everyday life, security and livelihood choices. It also brings attention to the human consequences of living in areas affected by long-term mining activities. Rawal, Bansal and Bansal (2016) studied land acquisition and income disparities in a village in Jharia, Jharkhand. Their work draws attention to the unequal economic consequences of land acquisition and raises questions about whether compensation and changes in employment adequately address the loss of traditional livelihoods. Noy (2022) examined the politics of dispossession and compensation in the eastern Indian coal belt. The study is useful in understanding the institutional and social dimensions of mining-related displacement, particularly the difficulties associated with compensation, land rights and the distribution of economic benefits. A more recent study, "A Just Transition or a Downward Spiral? Land and Livelihood Transitions to and Away from Coal Mining in India" (2025), discusses the challenges involved in moving away from coal-dependent economies. It reinforces the need to consider alternative employment opportunities and livelihood security when planning changes in the mining sector.

The literature suggests that mining should not be assessed only in terms of mineral output, industrial production or revenue generation. Its effects on land ownership, agriculture, employment patterns, environmental quality and community welfare must also be considered. Although these studies provide valuable insights into mining-related social and economic changes, there remains a need to bring agricultural concerns, rural development and the management of mining revenues into a comparative district-level discussion. The present study attempts to contribute to this discussion through the cases of Bokaro and Dhanbad.

2.1 Mining And Its Contribution To Jharkhand's Economy

Mining is one of the major economic activities in Jharkhand. The availability of mineral resources has encouraged the establishment of large industrial units and supported the growth of associated sectors. The contribution of mining to the state economy can be understood through the following areas.

2.1.1 Employment Generation

Mining provides direct employment to workers involved in extraction, transportation, processing and related operations. It also generates indirect employment through construction, trade, repair services, logistics and other

supporting activities. However, the employment opportunities created by large mining projects do not necessarily benefit all affected households. Local residents may lack the technical qualifications required for specialised positions, while traditional agricultural workers may find it difficult to shift to industrial occupations.

2.1.2 Industrial Development

Mineral resources have supported the growth of steel plants, thermal power stations and other manufacturing units. Bokaro, for example, has developed as an important industrial centre, with the steel industry playing a significant role in its economic identity. The availability of coal and iron ore has also encouraged the development of transport, energy and industrial infrastructure.

2.1.3 Infrastructure Development

Mining activities often lead to the construction and expansion of roads, railway networks, power facilities and communication infrastructure. Such development can improve connectivity and create opportunities for trade and other economic activities. Nevertheless, infrastructure development must be assessed in terms of who benefits from it. Industrial roads and transport facilities may support commercial activity without necessarily improving access to schools, healthcare facilities, drinking water or markets in nearby villages.

2.1.4 Government Revenue

Mining contributes to government revenue through royalties, taxes, auction proceeds and other statutory payments. A portion of the financial resources collected from mining-related activities can support the development of mining-affected areas through mechanisms such as the District Mineral Foundation. The effectiveness of these funds depends on their proper planning, utilisation and monitoring.

2.2 IMPACT OF MINING ON AGRICULTURE AND RURAL LIVELIHOODS

While mining contributes to industrial growth, its expansion can create significant challenges for rural communities. These challenges are particularly important in areas where agriculture and natural resources continue to support household income.

2.2.1 Loss of Agricultural Land

One of the most direct effects of mining is the conversion of agricultural land into mining sites, industrial areas, transportation corridors and waste-disposal zones. Land acquisition may reduce the area available for cultivation and affect the livelihood of farming households. Even when compensation is provided, the loss of agricultural land can have long-term consequences because land provides a recurring source of food and income. For small and marginal farmers, the loss of even a limited area of cultivable land may significantly affect household security.

2.2.2 Decline in Soil Quality

Mining operations involve excavation, removal of vegetation, movement of heavy machinery and the dumping of overburden. These activities can disturb the natural structure of the soil. Dust deposition and the improper disposal of mining waste may also affect the quality of nearby agricultural land. Where soil degradation occurs, farmers may experience lower productivity and increased dependence on agricultural inputs. The restoration of mined land is therefore essential for reducing long-term damage to agricultural livelihoods.

2.2.3 Water Pollution and Scarcity

Agriculture depends heavily on the availability of clean water. Mining can affect both surface water and groundwater through excavation, mineral processing, waste disposal and the movement of contaminated runoff. In mining-affected areas, changes in water quality may create difficulties for irrigation, livestock and domestic use.

Mining operations can also alter natural drainage patterns and affect the availability of water in nearby villages. The consequences are not limited to farmers. Households dependent on local wells, ponds and streams may also face difficulties in accessing safe water.

2.2.4 Air Pollution and Dust Emissions

Dust generated during drilling, blasting, transportation and mineral handling can spread to nearby agricultural fields and residential areas. Prolonged exposure to dust and particulate matter can affect human health and reduce the quality of the local environment. Dust deposition on vegetation may also interfere with plant growth, although the extent of the impact depends on local conditions and exposure levels.

2.2.5 Displacement and Loss of Traditional Livelihoods

Mining projects may require the acquisition of land and relocation of settlements. Displacement can affect not only agricultural landowners but also agricultural labourers, forest-dependent households and other groups whose livelihoods depend on local resources. Compensation is not always sufficient to replace the economic and social value of land. Families may also lose access to community networks, grazing areas, local markets and traditional sources of income. A livelihood-oriented rehabilitation policy must therefore extend beyond the payment of monetary compensation.

2.2.6 Health and Social Concerns

Mining-related dust, noise, contaminated water and changes in the local environment can create health risks for people living near mining sites. Health-related concerns may include respiratory problems, skin irritation and other illnesses associated with environmental exposure. The specific health outcomes depend on the type of mining, duration of exposure and local environmental conditions. Mining-affected communities therefore require access to regular health monitoring, safe drinking water and appropriate healthcare facilities.

2.2.7 The Jharia Coalfield: A Major Example

The Jharia coalfield in Dhanbad is one of the most significant examples of the long-term challenges associated with coal mining in Jharkhand. Underground coal fires, land subsidence, unsafe settlements and environmental degradation have affected parts of the region. The continuing problem of underground fires has created serious concerns relating to housing, health, land stability and the security of local residents. The Jharia experience demonstrates that mining-related development challenges do not necessarily end when extraction takes place. In some cases, their effects continue for decades and require sustained institutional intervention.

3. COMPARATIVE ANALYSIS OF TWO MINING DISTRICTS

The districts of Bokaro, Dhanbad represent different aspects of Jharkhand's mineral economy. Although the two have experienced the economic influence of mining, their local conditions and livelihood concerns differ.

3.1 Bokaro

Bokaro is widely recognised for its steel industry and its association with coal-based industrial development. Mining and industrial activities have contributed to employment generation, urbanisation and infrastructure expansion. However, rural settlements in and around industrial and mining areas may experience pressure on land and water resources. The expansion of industrial infrastructure can also change traditional patterns of land use. For rural households, the major concern is whether industrial growth generates sufficient local employment and improves access to basic services.

3.2 Dhanbad

Dhanbad is one of India's most prominent coal-mining regions. The Jharia coalfield is central to its mining economy. Coal extraction has supported industrial activity and employment, but the district also faces significant environmental and social challenges. Underground fires, land subsidence, dust pollution and displacement have affected communities in several mining-affected areas. Dhanbad illustrates the long-term consequences of mining when environmental management, land rehabilitation and community safety are not adequately addressed.

4. TOWARDS AN ERA OF SUSTAINABLE MINING

The growing environmental and social concerns associated with mining make it necessary to reconsider the conventional approach to mineral-based development. Mining is an important economic activity, and its complete elimination is neither a realistic nor a necessary development objective. However, mining practices must evolve to reduce environmental damage, protect local livelihoods and ensure that affected communities benefit from mineral resources.

The idea of sustainable mining involves balancing three important dimensions:

1. Economic viability
2. Environmental responsibility
3. Social welfare

A sustainable mining framework must consider the entire mining cycle, from project planning and extraction to land restoration and mine closure.

4.1 Restoration of Agricultural Land

Mining companies should be required to prepare and implement scientifically designed land-reclamation plans. Where mining activities have affected agricultural land, restoration should include soil treatment, land levelling, improvement of soil fertility and suitable plantation activities. The objective should not be limited to making abandoned mining land visually green. Restoration should aim to make land environmentally stable and, wherever feasible, suitable for productive use.

4.2 Afforestation and Ecological Restoration

Plantation activities should be planned according to local ecological conditions. Native species are generally more suitable for restoring local ecosystems than indiscriminate plantation of non-native species.

Mining companies should also undertake measures to control soil erosion, protect water bodies and restore disturbed habitats.

4.3 Scientific Waste Management

Mining generates overburden, tailings and other forms of waste. Improper disposal can damage land and contaminate nearby water resources. Scientific waste management requires proper storage, treatment, monitoring and safe disposal of mining waste. Where technically feasible, the reuse of waste materials should be encouraged.

4.4 Water Resource Management

Mining companies should adopt effective water-management systems to reduce contamination and minimise unnecessary consumption. Regular monitoring of groundwater and surface-water quality should be undertaken in mining-affected villages. Safe drinking water facilities and suitable water-treatment arrangements should be provided wherever required.

4.5 Responsible Mine Closure

Mine closure should not be treated as an activity to be undertaken only after extraction has ended. Closure planning should begin at the project-planning stage itself. A responsible closure plan should address land restoration, environmental monitoring, worker rehabilitation and alternative livelihood opportunities for affected communities.

5. CONCLUSION

Mining in Jharkhand must be made more sustainable and inclusive through effective land restoration, responsible water management, fair compensation and rehabilitation of affected communities. Transparent utilisation of DMF funds, local participation, skill development and alternative livelihood opportunities are equally important to reduce the adverse effects of mining on rural households. Mining companies and government agencies must also adopt proper mine-closure plans and regularly monitor environmental and social outcomes. Ultimately, the success of mining-led development should not be measured merely by industrial growth or mineral production, but by its contribution to the well-being of local communities. A balanced approach that integrates economic progress, environmental protection and rural livelihood security is essential for ensuring that Jharkhand's mineral wealth benefits both present and future generations.

REFERENCES

1. Dey, S., & Singh, R. (2021). Mapping precarious energy geographies: Exploring the lived experience of coal mining in Jharia, India. *Energy Research & Social Science, 82*, 102298. https://doi.org/10.1016/j.erss.2021.102298

2. IBM. (2014). *Manual on appraisal of mining plan*. Indian Bureau of Mines, Ministry of Mines, Government of India.

3. Ministry of Mines, Government of India. *District Mineral Foundation and Pradhan Mantri Khanij Kshetra Kalyan Yojana guidelines*. Government of India.

4. Noy, I. (2022). The politics of dispossession and compensation in the eastern Indian coal belt. *Critique of Anthropology*. https://doi.org/10.1177/0308275X221074831

5. Pelz, S., et al. (2024). The spatial and economic footprint of the coal industry on rural livelihoods in Jharkhand, India. *Energy Policy, 186*, 113973. https://doi.org/10.1016/j.enpol.2023.113973

6. Rawal, V., Bansal, V., & Bansal, P. (2016). *Land acquisition and income disparities: Case study of a village in Jharia, Jharkhand*. SSER Monograph 16/1.

7. Coal Controller Organisation. (2025). *RECLAIM and LIVES frameworks*. Government of India.

8. Government of Jharkhand. (2026). *Jharkhand Economic Survey 2025–26*. Finance Department, Government of Jharkhand.

9. *A just transition or a downward spiral? Land and livelihood transitions to and away from coal mining in India*. (2025). *Journal of Agrarian Change*. https://doi.org/10.1111/joac.70003

Impact of digital transformation on employee engagement, work–life balance and organisational resilience

Aman Aftab
Department of Masters of Business Administration,
Guru Gobind Singh Educational Society Technical Campus, Bokaro, Jharkhand, India.

Adnan Aslam
Department of Masters of Business Administration,
Guru Gobind Singh Educational Society Technical Campus, Bokaro, Jharkhand, India.

Shafaque Rahmat
Department of Masters of Business Administration,
Guru Gobind Singh Educational Society Technical Campus, Bokaro, Jharkhand, India.

Abstract:

Digital transformation has moved beyond the automation of isolated tasks to reshape organisational structures, work processes, employee experiences and the ways organisations respond to disruption. Digital technologies can improve connectivity, flexibility, access to information and decision-making, but may also intensify workloads, blur work–non-work boundaries, increase technostress and create uncertainty about employment and skills. This paper conceptually examines the impact of digital transformation on employee engagement, work–life balance and organisational resilience. Drawing on the Job Demands–Resources (JD–R) model, Social Exchange Theory and Dynamic Capabilities Theory, it develops an integrative framework in which digital transformation influences employee outcomes through two competing pathways. A resource pathway operates through digital competence, autonomy, collaboration, meaningful work and organisational support; a demand pathway operates through techno-overload, techno-invasion, role ambiguity, surveillance and job insecurity. Employee engagement and work–life balance are conceptualised as employee-level outcomes that may contribute to organisational resilience through learning, knowledge sharing and adaptive behaviour. Leadership support, digital skills development, employee participation, boundary-management practices and human-centred technology design condition these relationships. This is a conceptual paper based on an integrative review of scholarly literature; it does not report primary survey data or statistical testing.

Keywords: Digital transformation; employee engagement; work–life balance; organisational resilience; digital HRM; technostress; human-centred work design.

1. Introduction:

Digital transformation is increasingly understood as an organisation-wide process through which digital technologies alter business models, structures, processes, capabilities and stakeholder relationships. Cloud platforms, enterprise systems, artificial intelligence, data analytics, automation, collaboration applications and digital HR platforms have changed how work is allocated, monitored, coordinated and evaluated. These changes are not merely technological; they influence work quality, employee expectations, managerial practices and the capacity of organisations to anticipate and respond to uncertainty.

The employee experience of digital transformation is mixed. Digital tools can reduce repetitive work, improve access to information, support remote collaboration and provide greater flexibility. At the same time, constant connectivity, rapid technological change, digital monitoring and pressure to remain available can increase cognitive and emotional demands. Osmundsen (2024) highlights that digital transformation affects employees through changes in work design, emotional well-being and competency requirements, and stresses employee involvement, leadership, culture and training. A meta-analysis of 106 empirical studies reports both positive and negative associations between workplace digitalisation and employee outcomes, including engagement, performance, burnout and work anxiety (Xu et al., 2025).

The organisational implications extend beyond productivity. In volatile environments, organisations require the ability to anticipate threats, absorb disruptions, maintain essential operations and adapt. Digital capabilities can support these resilience processes through virtual access, collaboration and data-informed decisions (Wenzel et al., 2023). Yet technology alone cannot guarantee resilience: employees need the skills, motivation, well-being and trust required to use digital resources effectively. This paper examines the relationships among digital transformation, employee engagement, work–life balance and organisational resilience from a human-centred perspective.

2. Research Problem and Rationale:

A challenge in digital transformation is the tendency to evaluate success through technology adoption, process efficiency and financial performance while giving insufficient attention to employee experience. Employees may be expected to adapt to new platforms and processes without adequate training, consultation or workload adjustment. In such circumstances, digital transformation can increase rather than reduce work pressure.

Three outcomes deserve integrated consideration. Employee engagement reflects employees' energy, dedication and absorption in their work. Work–life balance concerns the extent to which individuals can manage work responsibilities alongside personal and family roles. Organisational resilience concerns the organisation's capacity to anticipate, cope with and adapt to disruptions. These outcomes are connected: engaged and supported employees may be more willing to learn, collaborate and contribute to change, while sustained work–life imbalance can weaken well-being and the human capacity required for adaptation.

Existing studies examine parts of this relationship, including digital transformation and employee engagement, digitalisation and employee well-being, and digital capabilities and organisational resilience. An integrated employee-centred framework connecting all three outcomes is therefore useful for theory development and management practice.

3. Objectives of the Study:

1. To examine the potential positive and negative effects of digital transformation on employee engagement.
2. To analyse how digital transformation may influence work–life balance through flexibility, connectivity and work intensification.
3. To explain the relationship between employee engagement, work–life balance and organisational resilience.
4. To develop an integrative conceptual framework identifying mechanisms and organisational conditions shaping these relationships.
5. To offer recommendations for human-centred and sustainable digital transformation.

4. Research Questions:

1. How does digital transformation influence employee engagement?
2. Under what conditions does digitalisation improve or undermine work–life balance?
3. How can employee engagement and work–life balance contribute to organisational resilience?
4. What organisational practices can reduce the negative employee consequences of digital transformation?

5. Review of Literature:

5.1 Digital transformation and the changing nature of work

Digital transformation involves the strategic integration of digital technologies into organisational activities and the associated redesign of processes, capabilities and relationships. It should be distinguished from digitisation, which converts analogue information into digital formats, and digitalisation, which uses digital tools to improve existing processes. Transformation implies more fundamental changes in how an organisation creates value and coordinates work. At employee level, these changes affect task content, skill requirements, autonomy, collaboration, performance evaluation and job security.

Osmundsen (2024) observes that the employee perspective has received less attention than organisation-level transformation and identifies work design, emotional well-being and competency requirements as important themes. A multi-level review of workplace digital transformation also emphasises skills, resilience, well-being, leadership, HR practices and organisational culture (Preparing Workplaces for Digital Transformation, 2021).

5.2 Digital transformation and employee engagement

Employee engagement is commonly associated with a positive, fulfilling work-related state characterised by vigour, dedication and absorption (Schaufeli et al., 2002). Digital technologies may strengthen engagement by improving access to resources, reducing routine effort, enabling timely feedback and facilitating collaboration. Digital learning platforms can support skill development and career growth, while analytics may help managers identify workload or development needs.

However, digitalisation can weaken engagement when employees experience continuous system changes, excessive monitoring, low control, unclear expectations or fear of technological replacement. Xu et al. (2025), in a meta-analysis of 106 empirical studies, report a double-edged relationship: digital technologies are associated with improvements in engagement and task performance, but can also intensify burnout, work anxiety and withdrawal-related behaviours. Liu et al. (2024) identify enabling and burdening pathways in the relationship between digital transformation and employee resilience, with thriving at work and job insecurity as important mechanisms.

The JD–R model offers an explanation. Digital autonomy, user-friendly systems, training, technical support and effective communication can function as job resources. Technological overload, complexity, constant availability and uncertainty can function as job demands. The effect of digital transformation therefore depends on the balance between resources and demands.

5.3 Digital transformation and work–life balance

Work–life balance refers to an individual's perceived ability to manage work and non-work responsibilities in a manner that is acceptable and sustainable. Digital tools can support balance through remote work, flexible scheduling, reduced commuting and more efficient coordination. These benefits are not automatic. When employees are expected to respond to messages outside working hours, attend virtual meetings across time zones or remain continuously connected, flexibility may become an extension of working time.

Technostress research identifies technology-related demands such as techno-overload, techno-invasion, techno-complexity, techno-insecurity and techno-uncertainty (Tarafdar et al., 2007). These pressures can blur boundaries and reduce recovery time. Mišíková and Jankelová (2024) examine how digital technologies, teleworking and corporate culture shape work–life balance and employee well-being. Work–life balance thus depends on organisational norms and managerial expectations, not merely on the availability of flexible technologies.

Clear communication norms, predictable availability expectations, reasonable workload allocation, employee choice over work arrangements and respect for non-working time can help preserve the benefits of digital flexibility.

5.4 Digital transformation and organisational resilience

Organisational resilience is the capacity to anticipate, prepare for, respond to and adapt to disruptions while maintaining or restoring essential functions. It includes learning and reconfiguration, not only recovery after a crisis. Digital technologies may support resilience through real-time information, distributed collaboration, digital service channels, scenario analysis and faster coordination.

Wenzel et al. (2023) explain how virtual access, virtual collaboration, data-driven decision-making, algorithmic reprogrammability and assisted decision-making can help firms adapt during crises. Research on digital transformation and organisational resilience in service SMEs also identifies a positive relationship between digital maturity and resilience-related outcomes (Building Organisational Resilience with Digital Transformation, 2022). Research on digital HRM reports that digital HR practices can strengthen anticipation, coping and adaptation capabilities (Digital Human Resource Management Influence on the Organisational Resilience, 2025).

Organisational resilience is not only a technological capability. Employees need competence and confidence to use digital systems, psychological resources to cope with change, and opportunities to share knowledge and solve problems. A digitally advanced organisation may remain fragile if employees are exhausted, excluded from decisions or dependent on systems without continuity plans.

5.5 Theoretical foundations:

5.5.1 Job Demands–Resources Model

The JD–R model proposes that job characteristics can be understood as demands requiring sustained effort and resources that help employees achieve goals, reduce demands or stimulate growth (Bakker & Demerouti, 2007). In

this paper, digital autonomy, training, technical support and collaborative platforms are potential job resources. Technological overload, complexity, constant connectivity and job insecurity are potential job demands. The model explains why the same technology may increase engagement in one setting but create strain in another.

5.5.2 Social Exchange Theory

Social Exchange Theory suggests that employees respond to perceived organisational support and fair treatment through reciprocal attitudes and behaviours (Blau, 1964). When organisations provide meaningful training, involve employees in technology decisions, protect privacy and respect work–life boundaries, employees may reciprocate through engagement, knowledge sharing and constructive participation in change.

5.5.3 Dynamic Capabilities Theory

Dynamic Capabilities Theory focuses on an organisation's ability to sense opportunities and threats, seize them and reconfigure resources in changing environments (Teece et al., 1997). Digital transformation can strengthen these capabilities through data, connectivity and flexible systems. Employee engagement, learning and well-being are important human foundations that help convert digital resources into organisational adaptation and resilience.

6. Conceptual Framework:

The proposed framework positions digital transformation as the focal organisational condition. Its influence is channelled through two parallel pathways. The enabling pathway includes digital resources, perceived autonomy, digital competence, collaboration and organisational support. It is expected to strengthen employee engagement and support work–life balance when technology reduces friction and provides genuine flexibility. The burdening pathway includes techno-overload, techno-invasion, techno-complexity, digital surveillance and job insecurity. It is expected to reduce engagement and undermine work–life balance.

Employee engagement and work–life balance are proposed as complementary employee-level contributors to organisational resilience. Engagement may support discretionary effort, learning, voice and knowledge sharing. Sustainable work–life balance may preserve employee energy, recovery and capacity to adapt. Digital competence, supportive leadership, employee participation and boundary-management policies are proposed as contextual conditions shaping the relationships.

Conceptual flow: Digital transformation → digital resources and digital demands → employee engagement and work–life balance → organisational resilience. Leadership support, digital competence, employee participation and boundary-management practices influence these relationships.

7. Research Propositions:

P1: Digital transformation that provides accessible, reliable and task-relevant digital resources is positively associated with employee engagement.
P2: Digital transformation that increases techno-overload, techno-invasion, techno-complexity or job insecurity is negatively associated with employee engagement.
P3: Digital transformation improves work–life balance when it increases employee control over work location, scheduling and task execution.
P4: The positive relationship between digital transformation and work–life balance is weakened when expectations of continuous availability and after-hours responsiveness are high.
P5: Digital competence strengthens the positive relationship between digital transformation and employee engagement.
P6: Perceived organisational support and employee participation strengthen the positive effects of digital transformation on engagement and work–life balance.
P7: Employee engagement is positively associated with organisational resilience through learning, knowledge sharing, proactive problem-solving and adaptive behaviour.
P8: Work–life balance is positively associated with organisational resilience by supporting employee well-being, recovery and sustained adaptive capacity.
P9: Employee engagement and work–life balance jointly contribute to the relationship between human-centred digital transformation and organisational resilience.

8. Discussion:

The framework suggests that digital transformation should not be treated as inherently beneficial. Outcomes depend on work design and implementation quality. When digital tools simplify tasks, improve information access and increase employee control, they can function as job resources. When they increase interruptions, monitoring, workload or uncertainty, they can function as job demands. This duality is consistent with evidence documenting both beneficial and adverse employee consequences of digitalisation (Xu et al., 2025; Liu et al., 2024).

The framework connects employee experience with organisational resilience. Resilience requires more than robust IT infrastructure: employees must interpret information, coordinate across boundaries, learn new processes and respond constructively under pressure. Engagement can encourage employees to contribute ideas and share knowledge, while work–life balance can help sustain the energy and well-being required for long-term adaptation. These are proposed conceptual pathways, not causal effects established by this paper.

Technology strategy, HR strategy and resilience planning should therefore be integrated. Digital implementation should be accompanied by employee consultation, skill development, workload review, privacy safeguards and clear boundaries around availability.

9. Managerial and HR Implications:

- Adopt a people-centred digital strategy and assess how new systems change tasks, workload, autonomy and employee experience before implementation.
- Invest in continuous digital learning through role-specific training, peer support and protected learning time.
- Redesign work rather than simply adding technology; remove redundant processes and ensure automation produces realistic workload reductions.
- Establish digital boundary norms covering response times, after-hours communication, virtual meetings and employee availability.
- Involve employees in technology selection, pilot testing, implementation feedback and evaluation.
- Train managers to recognise technostress, communicate change transparently and balance performance expectations with well-being.
- Use digital HR analytics responsibly, ensuring transparency, privacy, fairness and appropriate human oversight.
- Link digital capability with resilience planning through cross-training, knowledge documentation, continuity plans and scenario-based exercises.
- Monitor balanced indicators, including adoption, productivity, engagement, workload, work–life balance, learning, absence and resilience.

10. Proposed Methodology for Future Empirical Validation:

This paper is conceptual and does not claim to have collected primary data. The proposed framework can be tested in a future empirical study involving employees from manufacturing, services, education, healthcare or other technology-intensive sectors in India. A cross-sectional survey may be used initially, although a time-lagged or longitudinal design would provide stronger evidence about temporal relationships.

A structured questionnaire could measure perceived digital transformation, digital competence, employee engagement, work–life balance, technostress, perceived organisational support and organisational resilience. Established scales should be selected after checking validity, suitability for the Indian context and usage conditions. Sample size should be determined through an a priori power analysis based on the proposed model and analytical technique rather than selected arbitrarily.

Data analysis may include descriptive statistics, reliability and validity assessment, and regression-based or structural equation modelling. Researchers should assess common method bias, non-response bias and multicollinearity where relevant. Ethical safeguards should include informed consent, voluntary participation, anonymity and secure data handling. A time-lagged design could measure digital transformation at Time 1, engagement and work–life balance at Time 2, and resilience-related outcomes at Time 3.

11. Limitations and Future Research:

As a conceptual paper, this study does not provide empirical estimates or establish causal relationships. The propositions require testing across industries, organisational sizes and employee groups. Digital transformation is broad: AI-based automation, collaboration platforms, HR information systems and remote-work technologies may produce different effects. Future studies should distinguish technology types, intensity of use, implementation quality and employee control. Research may also examine differences by occupation, caregiving responsibilities, career stage, contract status and work arrangement.

Future work could investigate whether engagement and work–life balance mediate the relationship between digital transformation and resilience, and whether digital competence, leadership support, organisational culture and boundary-management practices moderate these relationships. Mixed-method research could explain why similar technologies produce different outcomes in different workplaces.

12. Conclusion:

Digital transformation is reshaping the contemporary workplace, but its success should not be judged solely by the speed of technology adoption or operational efficiency. It must also be evaluated by its effects on employees and the organisation's capacity to adapt sustainably. This paper proposes that digital transformation influences employee engagement and work–life balance through both enabling resources and intensified demands. These employee outcomes may contribute to organisational resilience by supporting learning, collaboration, knowledge sharing, well-being and adaptive action.

The central argument is that resilient digital organisations are not simply technology-enabled; they are people-enabled. Organisations should combine investment in digital infrastructure with investment in employee competence, supportive leadership, participation, ethical data practices and healthy work boundaries. A human-centred approach can help ensure that digital transformation creates value without compromising the engagement and well-being of the people who make organisational resilience possible.

References:

1. Bakker, A. B., & Demerouti, E. (2007). The Job Demands–Resources model: State of the art. Journal of Managerial Psychology, 22(3), 309–328. https://doi.org/10.1108/02683940710733115
2. Blau, P. M. (1964). Exchange and power in social life. Wiley.
3. Building organizational resilience with digital transformation. (2022). Journal of Service Management, 34(1), 147–171. https://doi.org/10.1108/JOSM-06-2021-0216
4. Digital human resource management influence on the organizational resilience. (2025). Organization Management Journal, 22(2), 111–125. https://doi.org/10.1108/OMJ-09-2024-2299
5. Liu, P., Zhang, F., Liu, Y., Liu, S., & Huo, C. (2024). Enabling or burdening? The double-edged sword impact of digital transformation on employee resilience. Computers in Human Behavior, 108220. https://doi.org/10.1016/j.chb.2024.108220
6. Mišíková, N., & Jankelová, N. (2024). Digital transformation and corporate culture: Shaping work-life balance and employee wellbeing – A systematic literature review. Industry 4.0, 9(6), 233–236. https://stumejournals.com/journals/i4/2024/6/233
7. Osmundsen, K. S. (2024). Digital transformation from an employee perspective: A literature review. Beta, 38(1), 1–18. https://doi.org/10.18261/beta.38.1.1
8. Preparing workplaces for digital transformation: An integrative review and framework of multi-level factors. (2021). Frontiers in Psychology. https://pmc.ncbi.nlm.nih.gov/articles/PMC8021873/
9. Schaufeli, W. B., Salanova, M., González-Romá, V., & Bakker, A. B. (2002). The measurement of engagement and burnout: A two sample confirmatory factor analytic approach. Journal of Happiness Studies, 3, 71–92. https://doi.org/10.1023/A:1015630930326
10. Tarafdar, M., Tu, Q., Ragu-Nathan, B. S., & Ragu-Nathan, T. S. (2007). The impact of technostress on role stress and productivity. Journal of Management Information Systems, 24(1), 301–328. https://doi.org/10.2753/MIS0742-1222240109
11. Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. Strategic Management Journal, 18(7), 509–533. https://doi.org/10.1002/(SICI)1097-0266(199708)18:7<509::AID-SMJ882>3.0.CO;2-Z

12. Wenzel, M., et al. (2023). Upgrading adaptation: How digital transformation promotes organizational resilience. Strategic Entrepreneurship Journal, 18(1), 128–164. https://doi.org/10.1002/sej.1483
13. Xu, G., Zheng, Z., Zhang, J., Sun, T., & Liu, G. (2025). Does digitalization benefit employees? A systematic meta-analysis of the digital technology–employee nexus in the workplace. Systems, 13(6), 409. https://doi.org/10.3390/systems13060409

Beyond Credit: Digital Financial Inclusion and AI Readiness among Women in Self Help Groups in Jharkhand

Uzma Aliya
BBA Department
Al Kabir Institute of Management and Technology, Mango, Jamshedpur
Email: Sheikhaliya2428@gmail.com

Abstract:

Financial inclusion depends on women's ability to use services and participate in financial decisions, as well as their access to accounts and credit. This paper examined a framework linking digital financial activity, practical support and willingness to use AI-assisted financial information among women in self-help groups in Jharkhand. Analysis of 200 simulated questionnaire records used descriptive statistics, reliability estimates and Spearman correlations. The digital-use index averaged 1.69 on a 0–4 scale; confidence, decision participation and AI willingness averaged 2.93, 2.95 and 2.96 on 1–5 scales. All eleven associations met the Bonferroni-adjusted significance criterion. These results demonstrate an analytical procedure and require validation with field data. The framework highlights the importance of distinguishing device access from independent operation and AI awareness from willingness to try a described service.

Keywords: self-help groups; digital finance; artificial intelligence; financial agency; Jharkhand.

1. Introduction:

Bank accounts and credit facilities create opportunities for financial participation, but their availability does not show whether women can use them confidently. A member may receive money digitally while relying on a relative to inspect the balance or complete a payment. Access, operational capability and influence over financial decisions therefore need separate assessment.

Self-help groups (SHGs) provide a setting in which savings, borrowing and collective learning can be examined together. Digital payments and electronic records add new activities to this setting. Their value depends partly on access to a usable phone, understandable instructions and assistance when difficulties arise. Transaction frequency alone cannot establish whether members control their money or authorise its use.

AI-assisted information introduces further questions about familiarity, language, trust and accuracy. Interest expressed after an explanation describes a response to a proposed service; it does not demonstrate actual adoption. This paper developed a compact framework for examining these distinctions in the intended Jharkhand setting.

2. Review of literature:

The Global Findex framework separates account ownership from the use of payments, savings and credit, providing a basis for measuring inclusion through several activities (Demirgüç-Kunt et al., 2022). The GSMA (2025) discussion of affordability, literacy and digital skills similarly supports assessing the circumstances in which a phone can be used. These perspectives informed the separation of device access, frequency of use and practical capability.

Kabeer (1999) treated empowerment as a relationship between resources, agency and achievements. Her approach helps explain why digital access cannot stand for meaningful choice. Research on Indian SHGs examined women's empowerment within a broader institutional setting (Bali Swain & Wallentin, 2009), while Brody et al. (2015) documented variation across programmes and contexts. Together, these studies supported retaining decision participation as a distinct outcome rather than inferring it from transactions.

Technology acceptance research distinguished perceived usefulness from perceived ease of use (Davis, 1989). This distinction informed the treatment of AI willingness, usefulness and confidence as related but separate measures. Glikson and Woolley (2020) also showed why trust in AI requires specific examination. The NIST (2023) framework identified reliability, transparency and privacy as considerations for evaluating systems; favourable user perceptions alone cannot establish those properties.

The present framework brought these strands together without claiming to validate an existing acceptance model or provide representative estimates for Jharkhand. It connected current financial activity with limited indicators of agency and responses to a proposed AI information service.

3. Research objectives:

- Describe smartphone access and current use of digital financial services.
- Examine relationships between digital use, confidence, support, perceived barriers and decision participation.
- Assess prior AI awareness and willingness expressed after a standard explanation.
- Identify assistance preferences and priorities for subsequent field investigation.

4. Research hypotheses:

Eleven two-sided hypotheses were examined. For each comparison reported in the results table, H_0 stated that the population monotonic association was zero ($\rho = 0$), while H_1 stated that it differed from zero ($\rho \neq 0$). $H_0$1–$H_0$7 concerned digital use in relation to smartphone availability, confidence, family support, data cost, language difficulty, fraud concern and privacy trust. $H_0$8 concerned digital use and decision participation. $H_0$9–$H_0$11 related AI willingness to prior awareness, perceived usefulness and AI privacy trust. These hypotheses formed a statistical demonstration rather than a preregistered field study.

5. Research methodology:

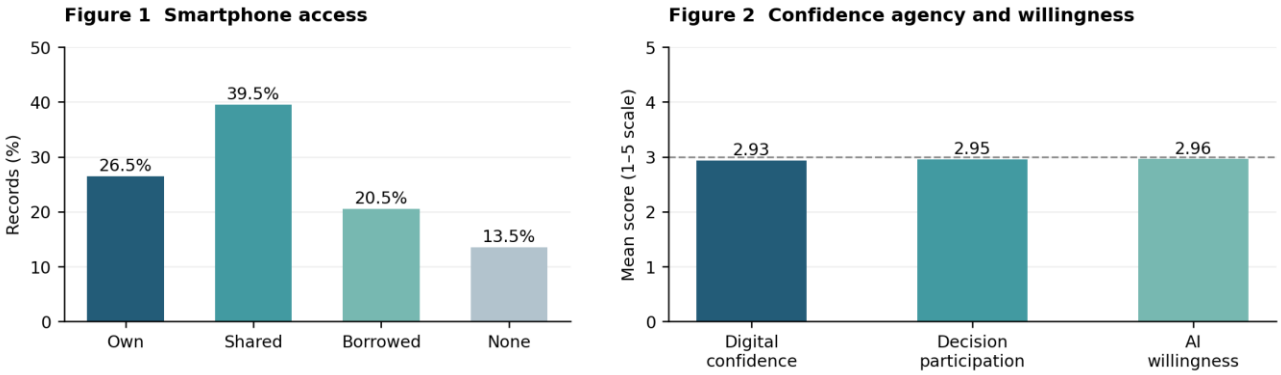
A quantitative analytical design was adopted using 200 simulated questionnaire records containing 76 variables. No participants were recruited and no field survey was conducted. Twenty illustrative group codes were used. The numerical relationships were influenced by a shared factor in the generation process; consequently, statistical significance demonstrated the analysis procedure rather than independent evidence about actual SHG members.

The 46-question instrument covered background characteristics, smartphone access, financial activities, confidence, support, concerns, decision participation and AI perceptions. Financial-use questions referred to the preceding three months in the proposed survey design. Frequency responses were coded 0–4, agreement statements 1–5 and decision participation 1–5. AI awareness preceded a standard explanation, while willingness was measured afterwards.

Digital use averaged making payments, receiving money and checking bank information (Q11–Q13). Personal use of SHG digital records was analysed separately because it depended on members' roles. Digital confidence averaged Q19–Q21, decision participation Q28–Q31 and AI willingness Q34–Q36. Records were checked for completeness, response ranges, coding consistency and skip rules; composite scores were recalculated without changing the underlying responses.

Descriptive statistics summarised the records. Cronbach's alpha assessed internal consistency for the three proposed reflective scales. Two-sided Spearman correlations examined the eleven associations. Bonferroni correction controlled the family of comparisons at $\alpha = .05$, equivalent to an unadjusted threshold of approximately .0045 per test. Sensitivity analyses excluded non-users or uncertain AI-awareness responses where relevant. No causal interpretation was applied.

6. Descriptive results:



Figures 1–2. Simulated records, N = 200. The dashed line marks the numerical midpoint; its meaning differs by scale. Digital use averaged 1.69 (SD 1.00; 0–4 scale). Reliability α was .839 for confidence, .835 for decision participation and .839 for AI willingness.

The largest age category was 40–49 years (30.0%); middle schooling accounted for 31.0%. Previous digital training was absent in 62.0%. Operation was independent in 28.0%, guided in 36.0% and delegated in 22.0%; 14.0% were non-users. Among 172 users, guided or delegated operation accounted for 67.4%.

7. Hypothesis testing and interpretation:

Null	Variable pair	Spearman ρ
$H_0$1	Smartphone availability and digital use	.773
$H_0$2	Digital confidence and digital use	.842
$H_0$3	Family support and digital use	.699
$H_0$4	Data-cost barrier and digital use	−.650
$H_0$5	Language barrier and digital use	−.686
$H_0$6	Fraud concern and digital use	−.681
$H_0$7	Privacy trust and digital use	.726
$H_0$8	Digital use and decision participation	.825
$H_0$9	Prior AI awareness and AI willingness	.621
$H_0$10	Perceived AI usefulness and AI willingness	.639
$H_0$11	AI privacy trust and AI willingness	.685

Note. N = 200. All eleven null hypotheses were rejected in the demonstration; every Bonferroni-adjusted p value was below .001. Coefficients were retained from the supplied analysis and reflect its generated-data structure.

Digital use was positively associated with enabling conditions and negatively associated with cost, language and fraud concerns. The association with decision participation remained positive after non-users were excluded ($n = 172$; $\rho = .744$; $p < .001$). Awareness and willingness also remained positively associated when uncertain awareness responses were excluded ($n = 180$; $\rho = .698$; $p < .001$). These checks did not overcome the limitations of the data-generation process.

8. Discussion:

Making and receiving digital payments occurred at least once in 83.0% and 84.0% of records, respectively, but weekly-or-more use occurred in only 28.0% and 28.5%. Personal use of SHG digital records was 52.5%. These differences show why occasional transactions, frequent use and participation in group bookkeeping should be reported separately. Low frequency also requires interpretation in relation to actual financial needs.

Prior AI awareness was recorded as yes in 49.0%, no in 41.0% and unsure in 10.0% of cases. Spoken interaction was preferred in 55.0%, while initial training followed by independent operation was preferred in 48.5%. Average willingness remained near the scale midpoint. Positive correlations therefore did not indicate uniformly high acceptance or demonstrated capability.

The most frequent barrier was other unspecified (48.5%), followed by unfamiliar language and no clear need (45.5% each). Other unspecified was also the largest support category (29.5%). Without explanatory responses, these selections cannot justify a precise intervention priority. Trusted assistance and local-language guidance require further needs assessment.

9. Principal findings:

- Shared phone access (39.5%) exceeded ownership (26.5%); 13.5% had no access. Guided or delegated operation accounted for 67.4% of users.
- Digital use averaged 1.69/4. Although 83.0% made digital payments, only 28.0% did so weekly or more often.
- Confidence (2.93), decision participation (2.95) and AI willingness (2.96) were near their midpoints. Digital use and decision participation correlated positively ($\rho = .825$).
- Spoken AI guidance was preferred by 55.0%; 48.5% preferred training before independent use. All eleven correlations met the adjusted criterion within the simulated data.

10. Suggestions:

- Pilot hands-on payment and balance-checking sessions: 62.0% lacked previous training and 67.4% of users relied on guided or delegated operation.
- Test local-language voice guidance with initial training, reflecting the 55.0% spoken-format preference and 48.5% preference for training before independent use.
- Include fraud prevention and answer verification; investigate unspecified barriers (48.5%) and support needs (29.5%) through follow-up questions before setting priorities.

11. Limitations and conclusion:

The records may not support population estimates or claims about actual SHG members. Ordinal averages, single-item measures and internally consistent scales required further validation. Future research should document sampling, consent, missing responses and group clustering, and complement perceptions with observed tasks. The framework's contribution was to separate access, use, decision participation and AI willingness, providing a clearer basis for subsequent field investigation.

References:

1. Bali Swain, R., & Wallentin, F. Y. (2009). Does microfinance empower women? Evidence from self-help groups in India. International Review of Applied Economics, 23(5), 541–556.
2. Brody, C., et al. (2015). Economic self-help group programs for improving women's empowerment: A systematic review. Campbell Systematic Reviews, 11(1), 1–182.
3. Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319–340.
4. Demirgüç-Kunt, A., Klapper, L., Singer, D., & Ansar, S. (2022). The Global Findex Database 2021. World Bank.
5. Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. Academy of Management Annals, 14(2), 627–660.
6. GSMA. (2025). The mobile gender gap report 2025.
7. Kabeer, N. (1999). Resources, agency, achievements: Reflections on the measurement of women's empowerment. Development and Change, 30(3), 435–464.
8. National Institute of Standards and Technology. (2023). Artificial intelligence risk management framework (AI RMF 1.0). https://doi.org/10.6028/NIST.AI.100-1

Waste to Wealth: Industrial Waste Reuse in Selected Manufacturing Units of Adityapur Industrial Area, Jharkhand

Sania Salim
BBA Department,
Al Kabir Institute of Management and Technology,
Mango, Jamshedpur
saniafatema1234@gmail.com

Abstract:

Industrial residues may retain economic value when they can replace other inputs or enter suitable recovery processes. This paper examines waste reuse with reference to 20 manufacturing units in the study setting of Adityapur Industrial Area, Jharkhand. It considers the types of waste generated, existing recovery options, operating benefits and constraints on implementation. An exploratory and descriptive approach guides the study, with structured questionnaires, semi-structured interviews and site observations forming the primary-data framework. Purposive selection focuses on production, environmental management and waste-handling functions. Corporate reports, government publications and technical guidance provide secondary evidence. Frequencies, percentages and material-flow ratios are used for the quantitative analysis, with Fisher's exact test and Holm's correction applied to three exploratory hypotheses. The analysis records an aggregate recovery rate of 65%, comprising 25% internal reuse and 40% external recovery. The estimated monthly operating surplus is ₹5.50 lakh within the specified accounting boundaries. Technology limitations and irregular demand are the most frequent barriers in the dataset. None of the tested associations remains statistically significant after adjustment for multiple comparisons. The paper supports cautious assessment of segregation, recovery costs, buyer coordination and material suitability, with conclusions restricted to the analytical scope of the study.

Keywords: Industrial waste; resource recovery; waste reuse; Adityapur Industrial Area; circular economy; Jharkhand.

1. Introduction:

Manufacturing produces residues alongside saleable goods. Although these materials are commonly treated as waste, some contain recoverable metals or other properties that make them useful in further production. Their value depends on whether they can meet the requirements of a receiving process at an acceptable cost. Converting waste into a resource therefore involves more than finding a buyer; it requires attention to quality, handling and the intended application.

Adityapur Industrial Area provides the geographical focus of this paper. The Jharkhand Industrial Area Development Authority identifies Adityapur as a major industrial hub within its administrative network (JIADA, n.d.). A study centred on this area allows waste reuse to be examined in relation to a defined industrial setting rather than through state-wide averages. The analytical categories cover steel and metal processing, fabrication and engineering, foundries, and construction-material manufacturing.

The paper considers how waste characteristics, segregation, processing capability and market access relate to recovery. It also examines the difference between gross receipts from waste sales and the operating benefit remaining after recovery expenses. Environmental gains are treated as matters requiring evidence: a commercial outlet does not, by itself, demonstrate that a material is safe or that its reuse reduces overall environmental impacts.

2. Review of literature:

The concept of industrial symbiosis helps explain how enterprises can use one another's residual materials. Chertow (2000) examines resource exchanges across organisational boundaries and the conditions through which such connections develop. This perspective is relevant because a by-product that has little use within one factory may satisfy an input requirement elsewhere. Establishing that connection requires compatibility between the material available and the receiving enterprise's technical and commercial needs.

Tata Steel's reporting provides an example of organised by-product utilisation. The company reports full solid-waste utilisation at its Jamshedpur, Kalinganagar, Meramandali and Gamharia locations for 2023–24, together with the development of processed slag products for construction uses (Tata Steel, 2024). These disclosures show that

recovery systems operate at industrial scale. They provide useful context, but the performance of a large integrated producer cannot be assumed to describe the circumstances of independent units in Adityapur.

Government evidence also points to substantial demand for selected industrial residues. According to the Ministry of Railways (2025), India generated 340.11 million tonnes of fly ash during 2024–25 and utilised 332.63 million tonnes. Roads and flyovers, cement production, and bricks and tiles were among the principal applications. These national figures establish the scale of utilisation but do not resolve the costs, quality requirements or transport conditions faced by a particular manufacturing unit.

Technical preparation is another recurring consideration. Guidelines issued by CSIR–Central Road Research Institute address the processing of steel slag into aggregates for road construction (CSIR–CRRI, 2024). Their relevance lies in distinguishing an untreated residue from a material prepared for a specified use. A recovery opportunity must consequently be assessed against the requirements of its intended application rather than inferred from the availability of waste alone.

The literature and institutional evidence connect resource recovery with cooperation, processing and dependable end uses. However, broad utilisation figures offer limited information about waste handling inside individual units. The present paper brings these considerations together through an analytical examination of waste quantities, segregation, buyer arrangements and recovery costs in the Adityapur study setting.

3. Research problem:

The research problem concerns the gap between generating a potentially recoverable material and establishing a workable reuse arrangement. Mixed storage, variable quality, limited processing facilities and uncertain demand may prevent a residue from becoming a useful input. The paper examines which materials warrant further recovery assessment and how operational and commercial conditions relate to the level of reuse.

4. Objectives of the study:

- To identify the principal waste streams and handling practices within the 20-unit analytical sample for Adityapur Industrial Area.
- To examine internal reuse and external recycling or recovery arrangements.
- To estimate operating benefits through recovery receipts, material savings and avoided disposal expenditure.
- To assess associations between waste reuse and segregation, regular buyer arrangements and transport-cost constraints.
- To identify practical barriers and offer modest suggestions relevant to manufacturing units in the study setting.

5. Research hypotheses:

Three pairs of hypotheses address the associations specified in the fourth objective. The remaining objectives are examined through descriptive and financial analysis.

H_{01}:Waste segregation and the level of waste reuse are independent.

H_{11}: Waste segregation and the level of waste reuse are associated.

H_{02}: A regular buyer or user arrangement and the level of waste reuse are independent.

H_{12}: A regular buyer or user arrangement and the level of waste reuse are associated.

H_{03}: Identifying transport cost as a major barrier and the level of waste reuse are independent.

H_{13}: Identifying transport cost as a major barrier and the level of waste reuse are associated.

6. Research methodology:

The research design is exploratory and descriptive, with Adityapur Industrial Area, Jharkhand, specified as the study location. The quantitative analysis uses a 20-unit dataset distributed across four manufacturing categories. The sampling framework is purposive: units are considered in relation to production activity, waste generation and willingness to provide information. The manufacturing unit, rather than each individual respondent, is the unit of analysis. The framework does not support statistical representation of the entire industrial area.

The primary-data collection framework combines one structured questionnaire per unit with semi-structured interviews and site observations, subject to organisational permission. Production managers, environmental personnel and waste-handling staff are the relevant information sources. The questionnaire covers waste type, monthly quantities, recovery routes, segregation, buyer access, transport constraints and operating benefits.

Weighbridge records, invoices and material registers are the appropriate sources for checking quantitative responses where access is permitted.
The interview guide addresses quality requirements, rejected consignments, payment terms, processing difficulties and coordination with waste users. Observation focuses on material identification, separation, storage, weighing and traceability. These methods are intended to complement the quantitative analysis. No interview quotations or site-specific observations are included without supporting records.
Secondary information is drawn from academic literature, official publications, corporate disclosures and technical guidance. These sources establish the wider context and help identify possible recovery routes. Their original reporting periods, organisational boundaries and definitions are retained; national or corporate totals are not combined with the 20-unit dataset.

7. Variables and analytical approach:

The quantitative information was summarised through counts, percentages and material-flow calculations. The monthly recovery rate was calculated as the quantity reused internally or directed to an external recovery use divided by waste generated during the reference month, multiplied by 100. The term reuse is used broadly in the analysis to include recycling and other documented recovery routes. Internal reuse and external recovery are shown separately. The calculation relates to current-month generation; legacy-stock movements require separate accounting.
For cross-tabulation, a recovery rate of at least 60% was classified as higher reuse, while a rate below 60% was classified as lower reuse. This cut-off is an analytical choice rather than an industry benchmark. Two-sided Fisher's exact tests were used because the comparisons involved small groups. Significance was assessed at 5%, followed by Holm's correction across the three tests. The tests describe associations within the dataset and do not establish causal effects.

8. Data analysis and interpretation:

Profile of units

Manufacturing category	Units	Share
Steel and metal processing	6	30%
Fabrication and engineering	6	30%
Foundry operations	4	20%
Construction materials	4	20%
Total	20	100%

Steel and metal processing and fabrication and engineering each accounted for six units. Together, they represented 60% of the sample; foundries and construction-material units contributed four units each. The analysis therefore placed greater weight on metal-related activities. This distribution describes the analytical sample, not the industrial composition of Adityapur.

Waste streams and recovery options

Waste stream	Units recording waste	Potential route
Metal scrap and offcuts	14	Metal recycling
Slag	8	Processed aggregate or other approved use
Ash	6	Suitable cement or brick inputs
Dust or sludge	7	Composition-specific recovery assessment
Packaging residues	12	Segregated material recycling

More than one waste category could be recorded for a unit, so the combined number of responses exceeded 20. Metal scrap and offcuts appeared in 14 units (70%), followed by packaging residues in 12 (60%), slag in eight (40%), dust or sludge in seven (35%) and ash in six (30%). The table identifies possible recovery directions rather than confirming material suitability. Dust, sludge and other composition-sensitive residues require further assessment before an application can be selected.

Waste generation and recovery

Monthly material flow	Quantity
Waste generated	2,000 tonnes
Internal reuse	500 tonnes
External transfer for recovery	800 tonnes
Total recovery	1,300 tonnes
Storage or disposal	700 tonnes

The monthly material balance recorded 2,000 tonnes of generation and 1,300 tonnes of recovery, giving an aggregate recovery rate of 65%. Of the quantity generated, 500 tonnes were allocated to internal reuse and 800 tonnes to external recovery, equivalent to 25% and 40%, respectively. The remaining 700 tonnes, or 35%, combined storage and disposal because separate destination quantities were unavailable. The aggregate percentage is based on total tonnage and is not the arithmetic mean of the individual units' recovery rates.

Practices and implementation barriers

Practice or barrier	Units	Share
Routine source segregation	10	50%
Regular buyer or user arrangement	12	60%
Higher reuse rate of at least 60%	10	50%
Transport cost as a major barrier	8	40%
Technology or processing constraint	11	55%
Irregular demand	9	45%
Storage constraint	8	40%

Processing or technology constraints were recorded for 11 units (55%), making them the most frequent barrier. Irregular demand followed at nine units (45%), while transport cost and storage constraints each appeared in eight units (40%). Routine segregation was recorded for half of the sample, and 12 units had regular buyer or user arrangements. The barrier categories allowed multiple responses. Their frequencies indicate prevalence within the dataset, rather than the monetary loss attributable to each constraint.

9. Economic assessment:

The financial assessment estimated the operating contribution of recovery. Receipts from external transfers were added to avoided raw-material purchases and disposal expenditure. Incremental segregation, processing, testing, transport and administrative expenses were then deducted. Internal material savings and external sales were treated as separate benefits to prevent the same quantity from being counted twice. The table presents a combined monthly calculation rather than a separate profit statement for every unit.

Monthly benefit or cost	Amount in rupees
External recovery revenue	6,00,000
Avoided material purchases	2,50,000
Avoided disposal costs	1,00,000
Total gross benefit	9,50,000
Incremental recovery costs	4,00,000
Net operating benefit	5,50,000

Gross monthly benefits amounted to ₹9.50 lakh, against incremental recovery expenditure of ₹4.00 lakh. The resulting operating surplus was ₹5.50 lakh. This figure excludes capital expenditure, borrowing costs, taxation and depreciation, and should not be interpreted as net profit. Sensitivity analysis reduced the surplus to ₹4.50 lakh when incremental costs increased by 25% with benefits unchanged. A 20% decline in gross benefits, with costs fixed, reduced the surplus to ₹3.60 lakh. These calculations show how the estimate responds to changes in costs and receipts.

10. Hypothesis assessment:

Comparison	Higher reuse	Lower reuse
Routine segregation	8	2
No routine segregation	2	8
Regular buyer arrangement	8	4
No regular buyer arrangement	2	6
Transport is a major barrier	2	6
Transport is not a major barrier	8	4

Each pair of rows compares the same 20 units on a different characteristic. Higher reuse was recorded in eight of the ten segregating units (80%) and two of the ten units without routine segregation (20%). Among units with a regular buyer arrangement, eight of 12 (66.7%) had higher reuse, compared with two of eight (25%) without an arrangement. Higher reuse occurred in two of eight units identifying transport cost as a major barrier (25%), against eight of 12 units without that reported barrier (66.7%). These are descriptive comparisons; the corresponding tests assess the strength of the statistical evidence.

Hypothesis	Odds ratio	Exact p value	Holm adjusted p	Decision after correction
H_{01}	16.00	0.0230	0.0690	Do not reject
H_{02}	6.00	0.1698	0.3396	Do not reject
H_{03}	0.167	0.1698	0.3396	Do not reject

11. Hypothesis testing:

The segregation comparison produced an odds ratio of 16.00 and an unadjusted two-sided p value of 0.0230. The regular-buyer comparison yielded an odds ratio of 6.00 and a p value of 0.1698. For the transport-barrier comparison, the odds ratio was 0.167 and the p value was also 0.1698. The latter two comparisons have the same p value because their cell counts reverse the exposure groups. An odds ratio below one indicates lower odds of higher reuse in the first-listed group.

Holm's procedure compared the smallest p value with 0.05/3, or approximately 0.0167. The observed value of 0.0230 exceeded this threshold, so none of the three null hypotheses was rejected after correction. Holm-adjusted p values were approximately 0.0690 for segregation and 0.3396 for both other comparisons. The descriptive contrasts therefore did not provide sufficient evidence for the alternative hypotheses after multiple-testing adjustment. Failure to reject a null hypothesis should not be read as proof that no association exists.

12. Framework for thematic analysis:

The qualitative framework groups the interview topics into four areas: material quality and segregation; processing and storage; buyer dependability and commercial terms; and transport and coordination. These categories guide the organisation of interview notes while allowing additional themes to be developed when the material supports them. Any thematic finding should remain traceable to a transcript or contemporaneous note, including accounts that differ from the dominant pattern.

Triangulation involves comparing questionnaire information with interview accounts, observable handling practices and available records. For instance, a statement about routine separation may be checked against storage arrangements and transfer documentation. This paper presents the framework without attributing quotations or field observations to named or anonymous participants.

13. Findings:

- Secondary publications identified established applications for selected steel-industry residues and fly ash, providing context for local recovery assessment.
- The 20-unit analytical sample recorded a 65% aggregate recovery rate, comprising 25% internal reuse and 40% external recovery.
- The monthly financial calculation yielded an operating surplus of ₹5.50 lakh within the stated cost boundaries.
- Technology limitations were the most frequent barrier in the dataset, followed by irregular demand.

- Higher reuse was more frequent in the segregation and regular-buyer groups and less frequent where transport cost was identified as a major constraint.
- None of the three associations met the 5% significance criterion after Holm's correction.

14. Suggestions:

- Units may keep straightforward records of the quantities and destinations of their main waste streams.
- Where practical, suitable recoverable materials may be separated and clearly identified at source.
- Technical advice may be obtained before a residue is introduced into an unfamiliar application.
- Processing and transport expenses may be reviewed alongside expected receipts or material savings.
- Regular communication with potential users may help clarify material quality, collection and payment arrangements.
- Periodic review of waste records may help units identify manageable improvements in recovery.

15. Limitations of the Study:

The analysis is limited to a 20-unit sample structure for Adityapur Industrial Area and cannot describe all enterprises in the area or across Jharkhand. Small group sizes reduce statistical precision, while differences between manufacturing activities restrict direct comparisons. The reliability of field-based estimates would depend on the completeness of responses and access to supporting records. The monthly aggregates do not show seasonal changes or separate storage from disposal. Laboratory characterisation and a full life-cycle assessment were outside the paper's scope, so material safety and net environmental gains were not established. Secondary publications also differ in definitions and coverage, limiting direct comparison with unit-level measures.

16. Conclusion:

Waste reuse offers a possible route to better material efficiency when recovery processes match the properties of the residue and the requirements of its user. The analysis for the Adityapur study setting illustrates how recovery quantities, operating costs and market arrangements can be examined together. Its financial estimate is positive within the stated boundaries, but the hypothesis tests do not support firm conclusions about the three associations after adjustment. Simple records, careful separation and informed assessment of costs and end uses are reasonable starting points for further evaluation. The paper does not establish that every waste stream is suitable for reuse or that recovery is profitable under all operating conditions.

References:

1. Chertow, M. R. (2000). Industrial symbiosis: Literature and taxonomy. Annual Review of Energy and the Environment, 25, 313–337. https://doi.org/10.1146/annurev.energy.25.1.313
2. CSIR–Central Road Research Institute. (2024). Guidelines for processing steel slag as processed steel slag aggregates for road construction. Ministry of Steel. https://steel.gov.in/sites/default/files/2025-02/Guidelines-for-Processing-Steel-Slag%20-%20Final%20Version%20-%20July%202024.pdf
3. Ministry of Railways. (2025, August 25). India generated 340 million tonnes of fly ash in 2024–25, successfully utilized over 332 million tonnes. Press Information Bureau.
4. Tata Steel. (2024). Natural capital. Integrated report and annual accounts 2023–24.
5. Jharkhand Industrial Area Development Authority. (n.d.). Adityapur. Retrieved November 14, 2026,

An Intelligent Edge-AI Framework for Low-Latency and Energy-Efficient Internet of Things Services

Chandan Kumar[1], **Sneha Kumari**[2], **Archita Kumari**[3], **Amit Paul**[4]

B.C.A Department, GGSESTC, Bokaro – 827013, Jharkhand, India

Correspondence Author's Email Id: ap9331196@gmail.com

Abstract

The Internet of Things (IoT) comprises sensors and devices that perpetually monitor and sense the physical environment. Transmitting each data point to a remote cloud incurs latency, wastes network bandwidth and exposes sensitive data. Edge computing is the means by which to mitigate this by processing sensing data locally, and with Edge AI, devices and neighboring gateways can optimize within device capacity through local decisions. This paper describes a conceptual model that unifies IoT sensors, microcontrollers equipped with TinyML, an edge gateway and cloud analytics. A task-allocation policy determines where in the stack each task should execute, based on factors such as priority, latency, energy, device capacity, network quality and privacy. Based on the literature; no prototype and no experiments are yet completed; it therefore describes architecture and expected behavior, rather than measured performance. A future prototype using ESP32-class hardware, Tensor Flow Lite for Microcontrollers and MQTT is specified.

Keywords: Internet of Things, Edge Computing, Edge AI, TinyML, Computation Offloading, Energy Efficiency, Adaptive Task Allocation.

1. Introduction

IoT are found in sectors like agriculture, health, transport, industrial applications and smart buildings. Sensors gather data and connected systems either detect a scenario or event and respond or check the data to monitor conditions. Data can be stored and managed on cloud platforms at scale, but depending on the application it may not be feasible to send all sensor data to a remote server.

Edge computing involves computing closer to the source of the data. Edge AI introduces machine-learning inference locally. This enables a device or an intermediate gateway to perform some level of data classification, anomaly detection or basic decision-making without having to communicate each data point back to the cloud. This may reduce communications, but it depends on the device, workload and network.

There is no single location for processing that is right for any task. A microcontroller is good for short, quick tasks but it has limited memory and processing. A gateway might be able to do those tasks but it has much more capability, whereas the cloud can do big data analysis and model training, so an adaptive system would identify the right place to do each.

1.1 Problem statement

The challenge is to choose which IoT task to run on the device, an edge or a cloud while striking a compromise on the response time, energy consumption, computing power, network condition and data privacy. Sending too much data can be inefficient if all processing is cloud-only. Running everything on the edge can be inefficient.

2. Literature Review

2.1 IoT and cloud-centric architectures

IoT sensors collect information, communication networks transmit the data, and computers generate the results. The data can be stored on the cloud and help us gain insight. But transmitting all the readings in each sensor again and again will consume more internet data and gives the system reliance on good internet connectivity.

2.2 Edge computing
Edge computing performs data processing near the device, not necessarily at a remote cloud. This can result in faster system response, lower network data traffic and greater retention of local data. It can also result in longer device battery life. But small devices have limited memory and processing power, so system design has to determine what workload the device can perform, and what needs to be offloaded to a local gateway or the cloud.

2.3 TinyML and embedded inference
TinyML enables small chips – microcontrollers – to run simple machine-learning models. This can help a device identify patterns or spot unusual measurements without having to send all measurements to the cloud. The model has to be small enough to fit on the device and able to operate within its memory, speed and energy limits.

2.4 Computation offloading and resource allocation
Offloading moves work away from the gateway or the cloud Shakarami et al. and Feng et al. examine techniques that represent a trade-off between delay, energy, and computing resources; placement is a multi-criterion decision.

2.5 Edge intelligence, security and trust
Research on trustworthy edge intelligence shows that security, reliability, transparency and sustainability are just as important, as performance. These factors should be built into the system from the start not tacked on after the system is already built. Designing with them in mind from the beginning makes a difference. It leads to systems that work better over time and can be trusted by users.

2.6 Communication protocol
MQTT is an OASIS maintained light weight publish subscribe protocol which can be used to pass sensor updates and control information between devices, gateways and services. MQTT must still be secured using encryption, authentication, authorization and appropriate QoS levels.

2.7 Comparative review and research gap
IoT cloud offerings, edge computing, and TinyML are well established in the literature. However, there is a need for a clear policy for small scale IoT systems which can coordinate distributed processing in devices, gateways, and cloud. This paper provides an architecture for this. It is not a proven optimization.

3. Proposed Intelligent Edge-AI Framework

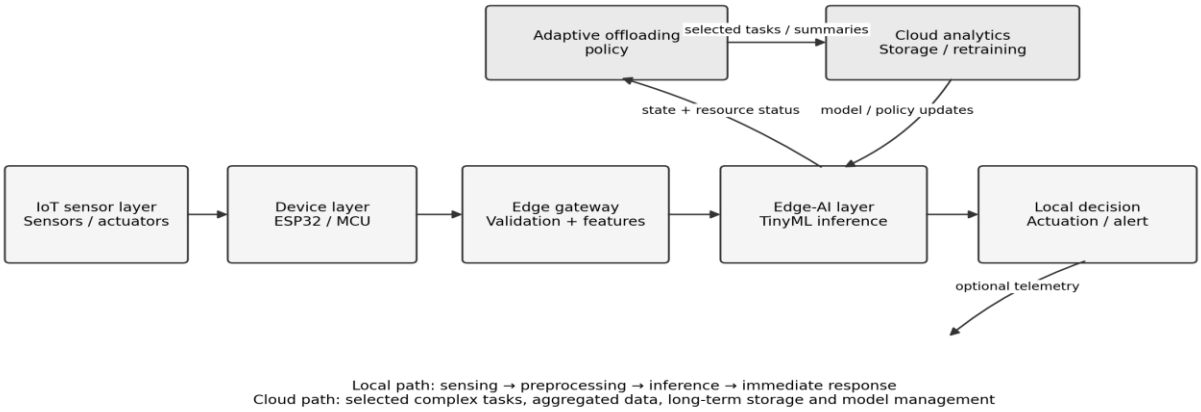


Figure 3.1: *Proposed Intelligent Edge-AI system architecture. The local path supports immediate inference and response; selected tasks and aggregated information may be forwarded to cloud services.*

3.1 IoT sensor and actuator layer
Sensors provide information about environmental factors like temperature, humidity, soil moisture, vibration or air quality. Actuators perform actions. Data should contain time stamps, device identification, units of measurement and some level of data validation to help identify any missing or unlikely data points.

3.2 Device layer: ESP32 or equivalent microcontroller
Sensors are used to gather information such as temperature, humidity, soil moisture, vibration or air quality. Actuators cause an action. The data should include timestamps, device information, units, and some form of data validation to help see if there are any missing or unlikely data points.

3.3 Edge-gateway processing layer
More processing and storage power is available at the gateway. This can be used to validate messages, aggregate data from multiple devices, process more complex models and manage task allocation. This could be done using a small Linux computer or an industrial gateway.

3.4 AI and TinyML inference layer
Lightweight models can be executed on the device to handle simple, time-critical tasks. Larger models can be executed at the gateway. Models should be trained on representative data, and validated for accuracy, memory requirements, speed and resilience. A rule-based fallback should be available for low confidence scenarios.

3.5 Adaptive communication and task-allocation layer
The task allocation layer takes into account the deadline, data size, computing load, and privacy level of the task, along with the device resources and network quality. It eliminates any allocations that do not satisfy the basic requirements (e.g., sending private data to the cloud). Then it chooses the most cost-effective allocation.

3.6 Cloud analytics layer
The cloud provides long-term storage, dashboards, analysis of historical data, training of models and controlled updates. Data should only be sent when required, preferably in aggregated form. Updates should be authenticated, version controlled and be rollable back.

4. Research Methodology and Working Process

This paper describes how the Edge-AI system could be implemented. We looked at published literature about IoT, Edge Computing, TinyML and Cloud Processing. From this we identified key principles that we would like to incorporate into our system. We then designed the system and determined what functionality each component would have. We described how the system would select to process tasks on the sensor, local gateway or Cloud. We then compared each of the three options based on their response time, energy consumption, network usage and privacy. As this is a design based study the benefits would need to be validated practically.

4.1 Choosing where a task should be processed
When a task is received, a decision will be made on whether to process it on the IoT device, the local gateway or the cloud. This decision will be made based on the urgency of the task, availability of memory and processing power on the device, availability of a connection and if the data is private. Simple and urgent tasks can be processed on the device. If it is a larger task then it can be sent to the gateway. If it requires more processing power or data from many devices it can be sent to the cloud. Only processing locations which can safely process the task will be chosen.

If the chosen location is not available then a safe local method should be used or a non urgent task should be stored. Important actions should be carried out locally when possible and not left waiting for the cloud. These rules have been suggested but need to be tested on a working prototype.

4.2 When data should be sent to the cloud
The cloud is used to complete tasks that are larger than can be done by the device and gateway, or when information needs to be accessed from multiple locations, or to retrain a machine learning model. This can only be done if the task can be completed after a delay caused by transmitting information. To reduce the amount of network usage and ensure privacy only the important features or brief summary can be sent.

4.3 Taking action after processing
This will result in the system taking some action. For example sending an alert, turning an actuator on/off, logging an anomaly or updating a dashboard. It is a good idea to log some basic information such as when the decision was taken and how confident the model was. If the action has any impact on safety then additional validation and override should be provided.

5. Results and Analytical Discussion

Limitations such as memory constraints, power consumption, unpredictable networking, model drift, privacy and device heterogeneity must be taken into consideration during real-world deployment. Using small models, local fallback, encrypted communications, device characterization and regular validation can help mitigate these issues. Inference controlled actuators should have safety constraints and manual override.

This comparison is not based on any physical experimentation but instead analysis of the proposed design and review. Therefore there are no real latency or power consumption figures given. The expected behavior must be tested on the target hardware.

5.1 Qualitative comparison of processing approaches

Cloud-based processing provides a scalable solution but relies on network connectivity which can cause latency and increased traffic. Edge-based processing can provide low latency and local processing but is limited by the resources of the device. Adaptive hybrid processing provides a balance between the two, processing urgent and viable tasks locally and more complex or less urgent tasks at the gateway or cloud. The best option will depend on the workload and circumstances.

5.2 How we will check system performance

We will check two things: how quickly the system responds and how much energy it uses.

Total response time = Processing time + Data transfer time + Waiting time

Processing time is the time taken to analyse the data. Data transfer time is the time taken to send and receive data. Waiting time is time spent waiting for a device, gateway or cloud service. Local processing can avoid waiting for a cloud reply.

The system's energy use can be considered in this simple way:

Total energy use = Energy used by sensors + Energy used for processing + Energy used for communication

Sensors use energy to collect readings, the device or gateway uses energy to process them, and communication uses energy to send data. Energy used by a pump or other actuator can also be included.

We can compare how many tasks are completed successfully for the energy used:

Simple efficiency comparison = Number of successful tasks / Energy used

A setup that completes the same tasks using less energy is more energy-efficient. A task is successful when its result is correct and arrives on time. These values will be measured in future testing; they have not been measured in this study.

To choose where a task should run, the system checks whether the device has enough capacity, whether the task can finish on time, whether the network is available, and whether privacy rules allow data sharing. Safety and privacy always come first.

6. Future Scope and Experimental Validation

For a prototype in the future, the ESP32, sensors, gateway, MQTT and cloud endpoint can all be used. An adaptive, cloud only and edge only system can be compared using the same inputs, model versions and testing under normal and poor network conditions, packet loss and disconnection. Response time, energy usage, network usage, model quality and resource usage can be measured. The hardware and software configurations can be recorded and tests repeated.

7. Conclusion

This paper describes a conceptual Edge-AI system which uses IoT sensors, ESP32-class devices, TinyML, Edge Gateways, dynamic task scheduling and Cloud analytics to address the problem of where to run tasks. It ensures urgent and possible tasks can be run on the device, whilst gateways allow greater workloads to be handled. The Cloud can still be used for long term storage, large-scale analytics and model management.

This paper provides a holistic, policy based design based on existing research. Practical solutions are presented alongside a plan for future testing. As no prototype or experimentation has been carried out, the benefits of this framework cannot be confirmed. The framework will need to be implemented and tested under controlled conditions to see how well it would perform with a particular IoT application.

References:

[1] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of Things (IoT): A vision, architectural elements, and future directions," Future Generation Computer Systems, vol. 29, no. 7, pp. 1645–1660, Sept. 2013, doi: 10.1016/j.future.2013.01.010.

[2] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, "Edge Computing: Vision and Challenges," IEEE Internet of Things Journal, vol. 3, no. 5, pp. 637–646, Oct. 2016, doi: 10.1109/JIOT.2016.2579198.

[3] N. Abbas, Y. Zhang, A. Taherkordi, and T. Skeie, "Mobile Edge Computing: A Survey," IEEE Internet of Things Journal, vol. 5, no. 1, pp. 450–465, Feb. 2018, doi: 10.1109/JIOT.2017.2750180.

[4] R. David et al., "TensorFlow Lite Micro: Embedded Machine Learning for TinyML Systems," Proceedings of Machine Learning and Systems, vol. 3, 2021. [Online]. Available: https://proceedings.mlsys.org/paper_files/paper/2021/hash/6c44dc73014d66ba49b28d483a8f8b0d-Abstract.html.

[5] A. Shakarami, M. Ghobaei-Arani, and A. Shahidinejad, "A survey on the computation offloading approaches in mobile edge computing: A machine learning-based perspective," Computer Networks, vol. 182, Art. no. 107496, Dec. 2020, doi: 10.1016/j.comnet.2020.107496.

[6] C. Feng, P. Han, X. Zhang, B. Yang, Y. Liu, and L. Guo, "Computation offloading in mobile edge computing networks: A survey," Journal of Network and Computer Applications, vol. 202, Art. no. 103366, June 2022, doi: 10.1016/j.jnca.2022.103366.

[7] X. Wang, B. Wang, Y. Wu, Z. Ning, S. Guo, and F. R. Yu, "A Survey on Trustworthy Edge Intelligence: From Security and Reliability to Transparency and Sustainability," IEEE Communications Surveys & Tutorials, vol. 27, no. 3, pp. 1729–1757, June 2025, doi: 10.1109/COMST.2024.3446585.

[8] OASIS, "MQTT Version 5.0," OASIS Standard, Mar. 2019. [Online]. Available: https://docs.oasis-open.org/mqtt/mqtt/v5.0/os/mqtt-v5.0-os.html.

Beyond code correction: a progressive hint-based ai tutoring framework for developing independent debugging and problem-solving skills across programming languages in technical education

Janvi Singh
B.C.A. Department,
GGSESTC, Bokaro

Anjali Kumari
B.C.A. Department,
GGSESTC, Bokaro

Amit Paul
B.C.A. Department,
GGSESTC, Bokaro

Abstract
The rapid introduction of generative artificial intelligence (AI) tools into programming education has made code generation, error detection, and automated troubleshooting widely accessible to students. However, direct access to AI code corrections can encourage passive copying, thereby reducing students' ability to track execution, diagnose logic errors, and develop transferable problem-solving skills. This challenge becomes even more apparent when students navigate programming languages that differ in syntax but share computational logic. This paper presents a conceptual advanced artificial intelligence teaching framework, initially focused on C, C++, and Python, with a language-independent design that aims to be extensible to any programming language in the future using modular language-specific adapters. Instead of immediately delivering corrected code, the system provides layered prompts (from broad concept prompts to targeted diagnostic questions) and adaptive prompts that fade in and out, becoming less useful as the learner's proficiency increases. A literature-based approach defines system components, learning rules, cross-language transfer exercises and subsequent assessment protocols. An initial black and white mockup of the website interface is provided; no requirement for a working software prototype or an empirical student dataset; contribution is a learning management structure that is consistent with outcomes-based education (OBE).

Keywords: Artificial Intelligence, Intelligent Tutoring System, Programming Education, Progressive Hinting, Adaptive Hint Fading, Multi-Language Programming, Debugging Skills, Computational Thinking.

1. Introduction

Learning programming with an engineering degree requires syntactic accuracy, mental tracing of a program, systematic error locating, and algorithmic problem solving.AI-powered generative programming assistants can generate code, interpret compiler diagnostics, and quickly repair defective programs. These opportunities lower barriers to entry, but their educational value depends on student engagement. When the wizard immediately produces the corrected program, the student can avoid the productive struggle necessary to understand why the original logic failed. Research on generative AI in computer science classrooms suggests that while AI explanations clarify unfamiliar constructs, unlimited access to comprehensive revisions can foster overconfidence and superficial understanding. Classroom implementations such as Code Aid have found that LLM wizards do not reveal complete solutions and are limited in providing a conceptual framework. This article addresses reliance on passive debugging through a structured workflow where students retain ownership of editing and validating code.

1.1 Research gap

Existing works have investigated data-driven hints in monolingual environments, and guardrail-based LLM helpers. Three drawbacks remain. Static cue systems may not be able to adapt explanatory details to repeated trials. Traditional LLM interfaces lack the ability for cues to automatically fade out as proficiency increases. And tutors often treat the languages separately, rather than reinforcing the shared algorithmic concepts of C, C++, and Python. This study proposes a unified architecture integrating progressive hinting, adaptive fading, and cross-language transfer.

1.2 Objectives and scope

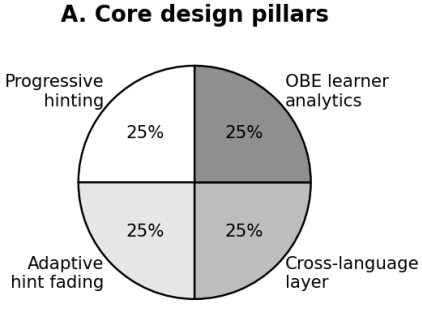
- Create a language-agnostic pipeline that analyzes student code and task specifications and provides spoiler-free step-by-step guidance.
- Adjust the specificity of indices as complexity increases and as independent debugging improves. We will initially support C, C++, and Python, but by separating language-specific syntax and diagnostics from general programming concepts, we will develop language-specific modular adapters that can support any programming language in the future.
- Identify future evaluation metrics: debugging accuracy, error localization time, queue dependencies, multilingual transfer.
- Define future evaluation measures: debugging accuracy, fault-localization time, hint dependency, and cross-language transfer.

Scope boundary: This is a conceptual and architectural study grounded in educational literature. A visual interface mock-up is included to illustrate the proposed student-facing experience; it is not a functioning software prototype. Neither complete student data nor empirical test results were reported.

2. Literature review

Intelligent tutoring systems (ITS) are intended to provide personalized instruction in programming labs. Rivers and Koedinger introduced ITAP, a Python tutor that generates personalized hints based on student answer status [4].Price et al.studied the problem of open data-driven programming [5].However, traditional AST-based tutoring requires significant language-specific development.

Large-scale language models make it possible to provide natural language descriptions in a variety of languages.Sun et al.studied students' programming behavior under the guidance of ChatGPT [1], and Jošt et al.discussed the impact of LLM on programming learning outcomes and the risks of unguided code generation [2].Kazemitabar et al.evaluated CodeAid, which uses conceptual guidance to preserve student agency [3].Grounded in the Zone of Proximal Development and cognitive load theory [6], the proposed framework combines multi-tier support with cross-language learning.



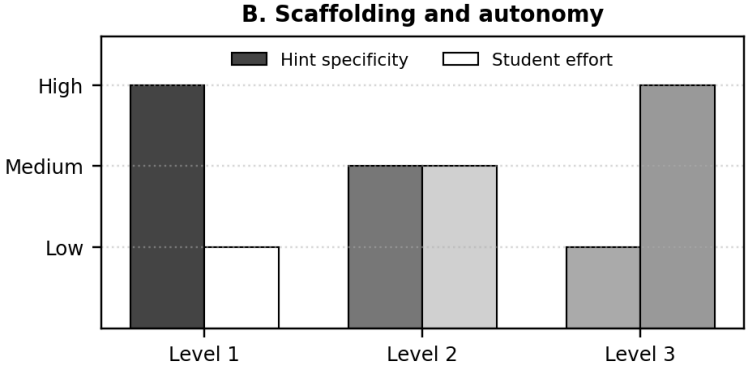


Figure 2.1. (A) *Four equally weighted architectural pillars (conceptual allocation). (B) Designed relationship between hint specificity and student effort; not empirical data.*

3. Proposed Framework Architecture

This framework acts as an educational intermediary between the learner and the AI inference engine. Instead of passing requests directly to code generation prompts, filter each response using prompt levels and student profile policies. It never overwrites the student's source file: learners interpret the diagnostic clue, modify their code, and resubmit it for verification.

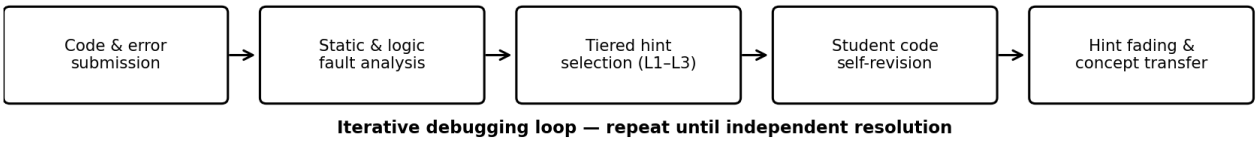


Figure 3.2: *End-to-end workflow governing fault analysis, hint selection, learner revision, and adaptive fading.*

3.1 Working principle of the model

1. Context Code Submission: Students choose C, C++, or Python and submit source code including the problem statement, expected output, and available compiler or runtime messages.
2. Multilevel error classification: The analyzer separates syntax errors (such as missing semicolons or indentation), runtime errors (such as segmentation errors and division by zero), and semantic logic errors (such as per-unit boundaries and uninitialized accumulators).Initial sending of tips. The hint manager starts at the lowest effective support level (level 1) to encourage self-diagnosis.
3. Check and resubmit to students: The student checks the concept, modifies the program independently, and submits another attempt. Adaptive escalation or decay: Repeated failures with the same error can trigger Level 2 or Level 3.Continued success will make future guidance more concise. Post-solution analysis and transfer: After passing the test, the feedback module can ask students to generalize the algorithmic principle and apply it in another supported language.

The cycle repeats until the learner resolves the issue independently. Complete reference solutions remain withheld by default and may be unlocked only under faculty policy after genuine attempts.

3.2 Progressive hint hierarchy and adaptive hint fading

To prevent cognitive overload while avoiding answer leakage, hints are organized into three bounded tiers. The example below concerns an off-by-one array traversal bug.

Scaffolding tier	Pedagogical objective	Illustrative hint (array boundary bug)	Information constraint
Level 1: General conceptual hint	Activate prior knowledge and direct attention to the relevant construct.	"Review how your loop controls the upper boundary when accessing elements in a fixed-size collection."	No line numbers, variable names, or syntax disclosed.
Level 2: Focused diagnostic hint	Localize the faulty logic block and highlight state transitions without providing code.	"Trace the value of index i on the final iteration of your for loop when i == n."	Identifies block/variable; repaired expression withheld.

Level 3: Targeted Socratic hint	Prompt edge-case verification so the learner constructs the exact fix.	"If an array has n elements indexed from 0 to n − 1, what happens when i <= n is evaluated?"	Guides mental tracing; student writes final operator.

Table 3.2.1: *Three-tier progressive hint specification with anti-spoiler constraints.*

Adaptive hint-fading rule: The Learner Progress Tracker maintains a sliding mastery score ($M^c \in [0,1]$) for each concept c, such as loops, pointers, or recursion. As a student repeatedly fixes bugs using only level 1 cheats, their proficiency increases and subsequent sessions remove the automatic level 2/3 escalation. If errors recur, supportive scaffolding is restored. Reference solutions remain restricted by faculty policy.

3.3 Multi-language learning & cross-language transfer

The Language and Concepts layer maps the Abstract Syntax Tree (AST) model of a particular language to a common standard concept graph across iterations, conditions, functions, arrays/lists, and memory/references. For example, after debugging pointer-based array inversion in C, students may encounter problems equivalent to one or two pointers in Python. The instructor will compare C's manual index boundaries to Python's indexing behavior. This design is intended to support future extensions to any programming language through language-specific adapters for parsing, compiler or interpreter diagnostics, syntax rules, and runtime feedback, while preserving a shared layer for common programming concepts. New languages can be integrated without changing the entire education system.

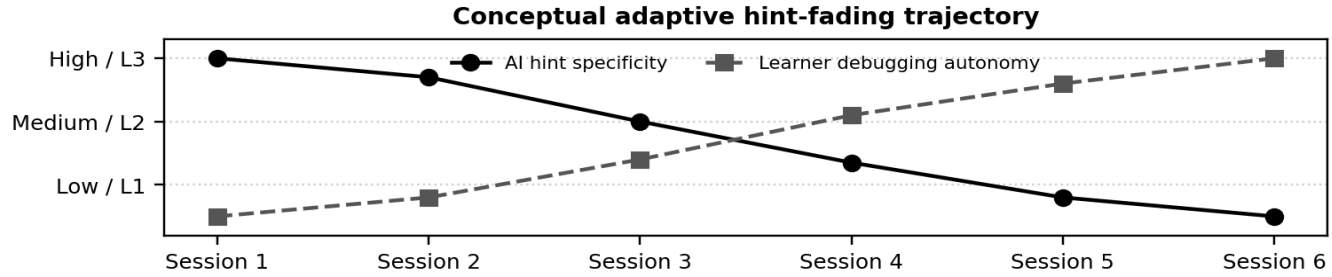


Figure 3.3.3: *Hypothesized inverse trajectory between AI hint specificity and learner debugging autonomy; requires empirical validation.*

4. Research Methodology

This investigation adopts a design-science and literature-grounded conceptual methodology. Prior studies on ITS, generative AI classroom guardrails and cognitive scaffolding inform the design requirements. Rather than claiming premature empirical results, this methodology defines modular components, educational rules for state transitions, and future experimental protocols.

4.1 Proposed system components

- Student Interaction Interface: A web-based code editor for C, C++, and Python that captures source code, task prompts, compiler/interpreter logs, and learner reflection without an "Auto-Fix" button.
- Code and Error Analyzer: Combines deterministic compiler/AST diagnostics with LLM semantic inspection to classify syntactic, runtime, and algorithmic faults.
- Hint Manager and Guardrail Filter: Selects Levels 1–3 and filters direct code patches before displaying guidance.

- Learner Progress Tracker: Logs attempts, time-on-task, hint levels, and self-correction success to update concept mastery and govern fading.
- Language and Concept Layer: Maps language-specific constructs onto shared concept taxonomy for cross-language exercises.
- Faculty Analytics and OBE Dashboard: Aggregates anonym zed class-level error patterns to help instructors target laboratory sessions and map outcomes to NBA/AICTE Course Outcomes.

4.2 Preliminary website interface mock-up

Figure 4.2.1 presents a proposed student-facing webpage for the tutoring framework. The interface combines a language selector and code editor with an error summary, three progressive hint levels, and a corrected-code reference. It illustrates the intended interaction flow; the controls and AI responses shown are illustrative and do not indicate a deployed or tested application.

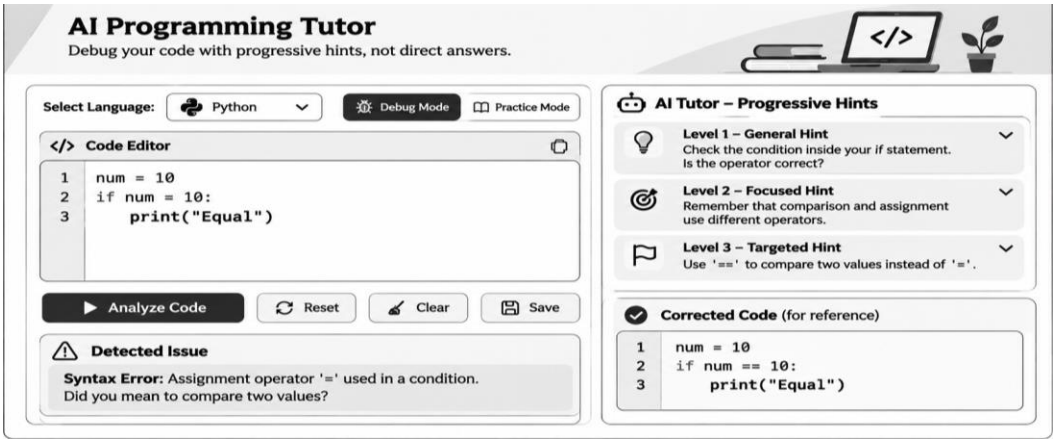


Figure 4.2.1: *Preliminary black-and-white website interface mock-up of the proposed AI Programming Tutor.*

4.3 Proposed experimental evaluation plan

After the prototype is implemented, it can be validated in an undergraduate programming lab and receive institutional ethical approval and informed consent. Introductory C/C++ and Python students can be divided into three groups using appropriate debugging tasks: (A) traditional direct AI assistant; (B) static one-level prompt helper; (C) Proposed progressive instant fading system. Pretests, delayed posttests without artificial intelligence, interactive telemetry, and student surveys will be analyzed. Evaluation data has not yet been collected.

5. Comparative analysis and expected outcomes

Table 5.1 contrasts the proposed framework with direct AI correction and basic static hints. This is a conceptual comparison of design expectations, not a report of measured findings.

Evaluation dimension	Direct AI code correction	Basic static hint support	Proposed progressive framework
Primary assistance	Immediately outputs corrected code or patch.	Provides fixed-depth clues regardless of learner history.	Delivers tiered hints (L1–L3) calibrated through fading.
Learner role	Passive reviewer; risk of uncritical copying.	Attempts fixes but may stall on vague clues.	Active diagnostician; hypothesizes, edits, traces, verifies.

Adaptability	Full solution depth from first prompt.	Similar hint detail across sessions.	Escalates during difficulty and fades as mastery rises.
Cross-language transfer	Focuses on isolated errors per query.	Often tied to single-language templates.	Links shared logic across C, C++, and Python.
Technical OBE alignment	May weaken authentic assessment of lab skills.	Encourages effort but lacks mastery analytics.	Builds independent debugging and tracks concept mastery.

Table 5.1*. Conceptual comparison of AI-assisted programming support paradigms; not empirical findings.*

5.1 Proposed quantitative and qualitative evaluation metrics

- Independent Debugging Accuracy (Adebug): Percentage of seeded syntax and logic bugs resolved correctly in a post-test without AI.
- Mean Time to Fault Localization (Tloc): Time from error presentation to the learner's first targeted edit.
- Hint Dependency Ratio (Rhint): Weighted ratio of Level 2 and Level 3 hints used per solved task; expected to decline over time.
- Conceptual Explanation Quality: Rubric-scored accuracy of the learner's explanation of why the original code failed.
- Cross-Language Transfer Score (Strans): Success rate when translating a debugged algorithm between C/C++ and Python without syntax hints.

5.2 Expected contributions to technical education

For institutions aligned with the ISTE mission, the framework offers three expected contributions: it can turn AI from an answer engine into a Socratic mentor; support multi-language fluency needed in software engineering; and provide faculty with diagnostic insight into student misconceptions without increasing grading workload.

5.3 Limitations and ethical safeguards

Because this study is conceptual, its effectiveness remains unverified until prototype deployment.LLM-generated hints may hallucinate or leak code patches, requiring automated guardrail checks and faculty review. Fault analysis without complete unit tests may misclassify subtle semantic errors. Hint-fading thresholds must be calibrated transparently so struggling learners are not penalized or unfairly labeled. Future classroom trials will require data privacy protection and consent.

6. Conclusion

Generative artificial intelligence is changing the way technology students interact with programming code, creating a need for learning systems that preserve independent reasoning. This paper presents a conceptual advanced prompt-based AI training system that replaces instant corrections with three layers of Socratic prompting, adaptive prompt fading, and concept transfer between languages using C, C++, and Python. Learners are still responsible for diagnosing and correcting their own programs, connecting the capabilities of AI with basic computational thinking.

Future work will proceed in phases: (1) develop and validate a modular prototype that initially supports C, C++, and Python, combining language-specific parsing and diagnostics with a guardrail-limited LLM prompt; (2) perform expert checks on the accuracy of hints and the reliability of anti-spoilers; and (3) evaluate the system in controlled classrooms. Beyond the initial implementation, the system will be designed to be extensible to any programming language through independent language adapters. Each adapter can provide associated parser, compiler, or interpreter interfaces, syntax rules, and diagnostic processing, while the shared guidance engine continues to provide progressive prompting, adaptive prompt fading, and concept-based guidance. This approach is not limited to Java or JavaScript; other languages can be integrated if appropriate adapters are developed and approved.

Research Status Declaration: This manuscript presents the conceptual architecture and preliminary model of the website interface derived from literature synthesis and instructional design. The interface shown is illustrative and not a working software prototype. No human-subject dataset or measured experimental result is reported; empirical claims are deferred until controlled validation.

References:

[1] D. Sun, A. Boudouaia, C. Zhu, and Y. Li, "Would ChatGPT-facilitated programming mode impact college students 'programming behaviors, performances, and perceptions? An empirical study," International Journal of Educational Technology in Higher Education, vol. 21, article 14, 2024. doi: 10.1186/s41239-024-00446-5.

[2] G. Jošt, V. Taneski, and S. Karakatič, "The Impact of Large Language Models on Programming Education and Student Learning Outcomes," Applied Sciences, vol. 14, no. 10, article 4115, 2024. doi: 10.3390/app14104115.

[3] M. Kazemitabaar et al., "CodeAid: Evaluating a Classroom Deployment of an LLM-based Programming Assistant that Balances Student and Educator Needs," Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24), article 650, 2024. doi: 10.1145/3613904.3642773.

[4] K. Rivers and K. R. Koedinger, "Data-Driven Hint Generation in Vast Solution Spaces: A Self-Improving Python Programming Tutor," International Journal of Artificial Intelligence in Education, vol. 27, no. 1, pp. 37–64, 2017. doi: 10.1007/s40593-015-0070-z.

[5] T. W. Price, Y. Dong, and T. Barnes, "Generating Data-driven Hints for Open-ended Programming," Proceedings of the 9th International Conference on Educational Data Mining (EDM), pp. 191–198, 2016.

[6] J. Sweller, P. Ayres, and S. Kalyuga, Cognitive Load Theory. New York, NY, USA: Springer, 2011. doi: 10.1007/978-1-4419-8126-4.

Impact of Artificial Intelligence on Digital Marketing for Small Businesses

Shourya Pratap
B.C.A. Department,
GGSESTC, Bokaro

Amit Paul
B.C.A. Department,
GGSESTC, Bokaro

Abstract
Artificial Intelligence (AI) is changing how businesses promote products, communicate with customers and make marketing decisions. For small businesses, digital marketing can be difficult because of limited budgets, fewer employees and strong competition. AI-based tools can support activities such as content creation, customer analysis, chatbots, personalized recommendations and campaign monitoring. This paper examines how these applications may improve marketing efficiency and customer engagement while considering cost, skills, privacy and reliability. The study follows a conceptual, literature-based approach and reviews published research on AI and marketing, together with material on digital adoption by small businesses. It proposes a practical framework in which customer data is collected responsibly, analyzed with suitable AI tools, converted into marketing actions and reviewed by a human decision-maker. No primary survey or experimental evaluation has been conducted; therefore, the benefits discussed are analytical expectations rather than measured outcomes. The paper concludes that AI can support small-business growth when tools are selected carefully and used with human oversight.

Keywords: Artificial Intelligence, Digital Marketing, Small Businesses, Marketing Automation, Customer Engagement, Personalization, Predictive Analytics.

1. Introduction

Digital marketing enables businesses to reach customers through websites, search engines, social media, email and online advertisements. These channels are useful to small businesses because they can promote products beyond their immediate local area and communicate directly with customers. However, effective digital marketing requires regular content, timely responses, audience understanding and continuous campaign monitoring.

Artificial Intelligence adds capabilities that can assist these activities. Machine-learning systems can identify patterns in customer data, natural-language tools can support written communication, and recommendation systems can help present relevant products. Generative AI can also help prepare captions, product descriptions and promotional ideas, although its output needs checking.

Small businesses may benefit because AI tools can reduce repetitive work and make some advanced marketing functions more accessible. However, benefits depend on data quality, staff skills, affordability and customer trust. AI should support rather than completely replace the judgment and personal relationships that are important to small-business marketing.

1.1 Problem statement

The problem addressed in this paper is how small businesses can use AI in digital marketing without creating unnecessary cost, privacy risks or dependence on automated decisions. They need a practical way to select suitable tools, use customer information responsibly and evaluate whether AI-supported marketing is producing useful results.

2. Literature Review

2.1 Artificial intelligence in marketing

Davenport et al. explain that AI can influence marketing strategy and customer behaviour through data analysis, automation and customer interaction [1]. They also emphasize that AI is most useful when it augments the work of human managers rather than simply replacing them.

2.2 Strategic use of AI
Huang and Rust present a strategic framework that considers AI in marketing research, marketing strategy and marketing actions [2]. Their work shows how AI can support market analysis, customer segmentation, targeting and personalized activities. This framework provides a useful basis for understanding how small businesses may apply AI across the marketing process.

2.3 AI and digital customer engagement
AI-powered chatbots and conversational systems can answer routine questions and provide information outside normal business hours. Personalization systems can use customer interests or past interactions to recommend relevant products. These applications may improve responsiveness, but they should provide a clear route to human assistance when a request is complex.

2.4 Generative AI and content creation
Generative AI can assist with drafting advertisements, social-media captions, product descriptions and email messages. It may reduce the time needed to prepare initial content, particularly for businesses with small marketing teams. The quality of the final material still depends on factual checking, brand knowledge and human editing [3].

2.5 AI analytics and decision-making
AI-supported analytics can organize large amounts of campaign and customer information, identify patterns and help estimate likely outcomes. Small businesses may use these insights to compare campaign performance, understand audience response and plan future promotions. Predictions should be treated as decision support, not as guaranteed results.

2.6 Adoption barriers, privacy and trust
Small businesses may face barriers such as limited budgets, lack of technical skills, uncertain data quality and difficulty selecting suitable tools. The OECD highlights that SMEs can experience constraints in resources and capabilities when adopting digital and AI technologies [4]. Marketing systems must also protect customer information and avoid unfair or misleading targeting.

2.7 Comparative review and research gap
The literature describes AI's potential in marketing automation, analysis, personalization and customer interaction. However, small businesses need practical guidance that connects these applications with limited resources and responsible data use. This paper addresses that need through an integrated conceptual framework. It does not claim to present a tested software product or a measured performance improvement.

3. Proposed AI-Enabled Digital Marketing Framework
The proposed framework connects six functions: customer and market data collection, data preparation, AI-supported analysis, marketing content and campaign planning, customer interaction, and performance review. Human supervision and privacy safeguards apply throughout the process. The framework is intended to be adaptable to small retailers, local service providers and online sellers.

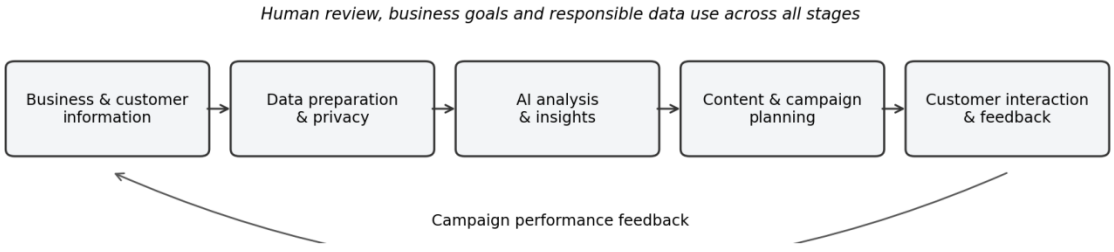


Figure 3. :. *Proposed AI-enabled digital marketing framework for small businesses. Customer and campaign information is analyzed to guide marketing actions; performance feedback informs later decisions.*

3.1 Customer and market data layer
The first layer gathers appropriate information from sources such as website visits, social-media engagement, customer enquiries, sales records and campaign reports. Businesses should collect only information needed for a clear purpose, keep records accurate and explain relevant data practices to customers.

3.2 Data preparation and privacy layer
Before analysis, information should be checked for missing, duplicate or incorrect entries. Personal details should be protected through access controls and suitable security practices. Where possible, aggregated or de-identified information can be used for general campaign analysis.

3.3 AI analysis and customer understanding layer
AI tools can help identify customer interests, frequently asked questions, product preferences and patterns in campaign response. A small business can use these insights to understand which products attract attention and which customer groups may be interested. The business should check whether the available data is representative before acting on a recommendation.

3.4 Content and campaign generation layer
Generative AI can prepare first drafts of social-media posts, product descriptions, email messages and advertisement ideas. Marketing staff or the owner should review the wording, prices, product claims, images and tone before publication. This review helps prevent inaccurate claims and keeps communication consistent with the business identity.

3.5 Customer interaction and personalization layer
Chatbots can handle routine enquiries, while recommendation tools can suggest products based on relevant customer interests. Automated messages should be useful, timely and easy to opt out of. Complex complaints, sensitive questions and unusual requests should be transferred to a human representative.

3.6 Campaign monitoring and feedback layer
The final layer tracks indicators such as reach, clicks, enquiries, conversion rate, repeat purchases and campaign cost. The business can compare these measures with its goals and adjust content, audience selection or spending. Regular review prevents automated campaigns from continuing when they are ineffective or inappropriate.

4. Research Methodology and Working Process
This paper uses a conceptual and literature-based research approach. First, published studies on AI in marketing were reviewed to identify common applications and strategic roles. Next, research on digital adoption by small and medium-sized enterprises was considered to understand resource and implementation barriers. The findings were organized into a practical framework covering data collection, analysis, content creation, customer interaction and performance review. The approach is descriptive and analytical; it does not involve a primary questionnaire, field survey or experiment.

4.1 Selecting suitable AI tools
A small business should begin by identifying a specific marketing problem, such as slow responses to customer enquiries, irregular social-media content or difficulty understanding campaign results. It should then compare tools based on cost, ease of use, language support, data handling, integration and the amount of human supervision required. A simple tool that solves a real problem may be more useful than a complex platform with features the business does not need.

If a tool produces inaccurate content, unsuitable recommendations or unexpected customer responses, the business should pause the activity and review the settings or workflow. Important customer issues should remain accessible to a human employee. These are proposed implementation practices and should be adapted to the size and nature of each business.

4.2 Responsible use of customer data
AI-based marketing should follow a clear purpose for collecting and using customer information. Businesses should limit access to data, use secure platforms, review privacy settings and avoid uploading confidential customer records

to tools that are not approved for such use. Customers should receive clear information about automated communication and available choices.

4.3 Evaluating marketing outcomes

After a campaign, the business should compare its results with the original objective. For example, a campaign intended to increase enquiries should be assessed using enquiry volume and quality, not only the number of views. Human review is needed to interpret results and decide whether a change is genuinely useful for customers and the business.

5. Results and Analytical Discussion

The literature-based analysis indicates that AI can support several digital marketing tasks, particularly content preparation, routine customer service, audience analysis and campaign monitoring. These applications may be valuable where staff time and marketing resources are limited. Their usefulness nevertheless depends on the quality of information, the suitability of the selected tool and the business's ability to review its output.

The discussion is analytical and based on the reviewed literature and proposed framework. No primary survey, business trial or controlled experiment was conducted. Therefore, this paper does not report measured increases in sales, engagement, customer retention or cost savings. Such outcomes would need to be evaluated in future field studies.

5.1 Qualitative comparison of traditional and AI-supported marketing

Traditional digital marketing can be managed manually, giving the business direct control but requiring time for content preparation, customer responses and performance analysis. AI-supported marketing can speed up selected tasks and help identify patterns, but it introduces costs, privacy considerations and a need for quality checks. A balanced approach combines automation for suitable routine tasks with human control over strategy and customer relationships.

5.2 How marketing performance can be evaluated

The business can evaluate AI-supported marketing by comparing suitable indicators before and after implementation.

$$\text{Engagement rate} = (\text{Total engagements} / \text{Total reach}) \times 100$$

Engagement includes interactions such as comments, shares, saves or clicks, depending on the platform and the campaign objective. The same definition should be used when comparing campaigns.

A business can also monitor the cost of acquiring a customer:

$$\text{Customer acquisition cost} = \text{Total marketing and sales cost} / \text{Number of new customers acquired}$$

This measure estimates the average cost of gaining a new customer during a defined period. It should be interpreted alongside customer quality, repeat purchases and the value of sales.

For campaigns intended to generate sales, conversion rate is another useful measure:

$$\text{Conversion rate} = (\text{Number of desired actions} / \text{Number of relevant visitors}) \times 100$$

The desired action may be a purchase, enquiry, booking or sign-up. These measures can help a business decide whether a campaign is meeting its goal, but they do not by themselves prove that AI caused the result. Other factors, such as price, season, product availability and market conditions, should also be considered.

A fair evaluation should compare similar campaigns, record the time period and account for changes in budget or audience. Customer privacy and the accuracy of automated content should be reviewed alongside performance indicators.

6. Future Scope and Research Directions

Future research can test the proposed framework with small businesses from different sectors, such as retail, food services and local professional services. A field study could compare selected AI-supported tasks with existing practices over a defined period. Useful measures may include response time, content production time, campaign

cost, engagement, conversion and customer satisfaction. Research should also examine language accessibility, affordability, data protection and the effect of human review on the quality of AI-generated marketing.

7. Conclusion

This paper examines the impact of Artificial Intelligence on digital marketing for small businesses and presents a conceptual framework for its responsible use. AI can assist with customer analysis, content preparation, chatbots, personalization and campaign monitoring. These capabilities may reduce repetitive work and help businesses make more informed marketing decisions.

The main contribution is a practical view of how small businesses can connect AI tools with clear marketing objectives, privacy safeguards and human oversight. AI should not be treated as a guaranteed route to higher sales or as a replacement for customer relationships. Its value depends on selecting suitable tools, checking outputs and evaluating results. Future field studies are needed to measure its actual impact across different types of small businesses.

References :

[1] T. Davenport, A. Guha, D. Grewal, and T. Bressgott, "How artificial intelligence will change the future of marketing," Journal of the Academy of Marketing Science, vol. 48, no. 1, pp. 24–42, 2020, doi: 10.1007/s11747-019-00696-0.

[2] M.-H. Huang and R. T. Rust, "A strategic framework for artificial intelligence in marketing," Journal of the Academy of Marketing Science, vol. 49, pp. 30–50, 2021, doi: 10.1007/s11747-020-00749-9.

[3] D. Grewal, C. B. Satornino, T. Davenport, and A. Guha, "How generative AI is shaping the future of marketing," Journal of the Academy of Marketing Science, vol. 53, 2025, doi: 10.1007/s11747-024-01064-3.

[4] OECD, "AI Adoption by Small and Medium-Sized Enterprises: OECD Discussion Paper for the G7," OECD Publishing, 2025, doi: 10.1787/426399c1-en.

[5] OECD, "The Digital Transformation of SMEs," OECD Publishing, 2021, doi: 10.1787/bdb9256a-en.

[6] OECD, "Generative AI and the SME Workforce: New Survey Evidence," OECD Publishing, 2025, doi: 10.1787/2d08b99d-en.

[7] P. Kotler and K. L. Keller, Marketing Management. Pearson Education.

InterviewLM: A shareable, retrieval-grounded conversational resume and mock interview platform

Harsh Kumar [1], **Sonu Kumar** [2], **Amit Paul**[3]

B.C.A Department, GGSESTC, Bokaro – 827013, Jharkhand, India

Correspondence Author's Email Id: ap9331196@gmail.com

Abstract

A conventional PDF resume presents a candidate's experience as a static document, leaving recruiters to search manually and candidates to prepare for interviews using generic question sets. This paper proposes InterviewLM, a shareable conversational resume platform that converts a candidate-approved PDF into a document-grounded chat experience. The system extracts resume text, organizes it into sections, creates searchable representations, and uses Retrieval-Augmented Generation (RAG) to answer questions from relevant resume passages. Recruiters can explore a candidate's skills, projects and experience through a shareable link, with responses accompanied by source sections or page references. Candidates can use a separate private mode to practise resume-specific mock interviews, receive follow-up questions and review structured feedback. The proposed design emphasizes evidence-based answers, explicit sharing controls, privacy safeguards and clear handling of missing information. InterviewLM is presented as a proposed system; its architecture and evaluation plan are described without claiming completed implementation or measured results.

Keywords: InterviewLM, Retrieval-Augmented Generation, Conversational Resume, PDF Parsing, Large Language Models, Mock Interviews, Vector Search, Explainable AI, Data Privacy.

1. Introduction

Resumes are commonly shared as PDF documents because the format preserves layout and is easy to distribute. However, a resume remains a mostly static representation of a person's education, skills, projects and work history. A recruiter who wants to understand a particular project, technology or contribution must locate the relevant text and may need to contact the candidate for clarification. Candidates face a related problem during preparation: generic interview question lists do not always reflect the specific experiences and claims presented in their own resumes.

Large language models (LLMs) can support natural-language interaction, but a model answering from general training alone may introduce details that are not present in a candidate's document. Retrieval-Augmented Generation (RAG) offers a way to connect generated answers with external evidence retrieved at query time [1]. In a resume setting, the source document can act as a small, candidate-specific knowledge base. Evaluation of RAG systems must consider retrieval relevance as well as the faithfulness and usefulness of generated answers [2].

InterviewLM is proposed as a two-mode platform built around one candidate-approved resume. In Recruiter Chat mode, a person with the share link can ask questions about the information contained in the resume and inspect the supporting excerpts. In Candidate Practice mode, the owner can rehearse technical, project-based and behavioral questions generated from their own experience, optionally using a supplied job description. The system is intended to make a resume easier to explore and interview preparation more personal, rather than to make automated hiring decisions.

The objectives are: (i) convert a static PDF resume into structured, searchable content; (ii) answer natural-language questions with retrieved evidence; (iii) provide a controlled shareable view for recruiters; (iv)

generate personalized mock interviews and actionable feedback; and (v) evaluate answer grounding, usability, privacy and system performance.

2. Literature Review

Lewis et al. introduced Retrieval-Augmented Generation (RAG), which combines a language model with a retriever over an external knowledge source [1]. This approach suits resume chat because answers can be based on passages retrieved from the candidate's own document. A typical pipeline separates document preparation, retrieval and answer generation, while retaining source metadata so responses can point back to the relevant resume section and page.

RAG evaluation research emphasizes that fluent output alone is not sufficient. Retrieval relevance, answer correctness, faithfulness to evidence and end-to-end usefulness are separate concerns [2]. In resume chat, a plausible answer may still overstate a role or combine details from different projects. InterviewLM therefore proposes source citations, a clear "not stated in this resume" response and candidate review of extracted text.

Interview preparation benefits from questions tied to job-related competencies and past behaviour. The U.S. Office of Personnel Management recommends clear, relevant questions and describes the STAR structure (Situation, Task, Action and Result) for eliciting examples [5]. NIST guidance also highlights the need to manage generative-AI risks across system design and use [3]. These ideas inform InterviewLM's combination of evidence-grounded recruiter chat and private, resume-specific practice.

3. Problem Statement and Research Gap

A PDF resume is easy to share but difficult to explore interactively. Recruiters may overlook details or repeatedly ask questions already answered in the document. Generic chatbots are not automatically grounded in a particular resume, while generic mock interview tools may ask questions unrelated to a candidate's projects. Parsing errors and unsupported AI responses can further reduce trust.

This proposal addresses the need for a unified, candidate-controlled workflow that makes resume information conversational while preserving links to the original PDF. Its focus is not only summarization or question generation: it retrieves evidence for recruiter queries, shows where answers came from, and reuses the approved profile for private practice without exposing interview responses.

4. Proposed System and Methodology

4.1 System Overview

InterviewLM has two principal roles: candidate and recruiter or resume viewer. The candidate uploads a PDF, reviews extracted content and creates a share link with chosen permissions. Recruiters use a read-only chat to ask about the shared resume. The candidate's private workspace stores mock interviews and feedback. A web interface can connect to backend services for parsing, retrieval, model requests and access control.

4.2 PDF Ingestion and RAG Workflow

After upload, a document processor extracts text while retaining page boundaries and layout cues. Scanned pages may need Optical Character Recognition (OCR), and uncertain text is flagged for candidate review. Content is organized into sections such as summary, skills, education, experience, projects and certifications. Section-aware chunks carry metadata such as page number and heading, allowing retrieved passages to be traced to the PDF.

Chunks are converted into vector embedding's and stored in a searchable index. For each question, the system retrieves relevant passages; keyword search or reranking may supplement vector similarity. The retrieved evidence and question are sent to the LLM with instructions to answer only from resume context, cite source pages or sections, and say when information is absent. This implements the retrieve-then-generate principle of RAG [1].

4.3 Recruiter Chat and Candidate Practice

Recruiter Chat supports questions such as "Which project used Python?" or "Where is cloud experience mentioned?". Answers include concise evidence excerpts and source locations. Share links may have expiry, access-code and revocation controls. The recruiter view exposes only candidate-approved resume content; private practice responses, notes and scores remain inaccessible.

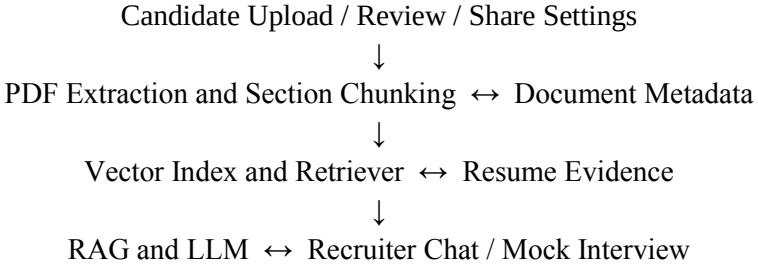
Candidate Practice uses the same resume index in a private workspace. The candidate selects a general, technical, project or behavioral interview and may add a job description. The system asks one question at a time, follows up on stated experience and gives feedback on relevance, clarity, structure and detail. A STAR prompt can help organize an example [5]. Feedback supports learning; it must not invent achievements or encourage unsupported claims.

4.4 Privacy, Security and Responsible Use

Resumes contain personal information, so processing and sharing require clear candidate consent. The design includes encrypted transfer and storage, access-controlled records, expiring or revocable links, deletion controls and defined retention periods. Documents are not placed in public indexes or used to train unrelated models without separate permission. Deployment should be reviewed against applicable data-protection requirements, including India's Digital Personal Data Protection Act, 2023 [4].

4.5 Proposed Architecture

The architecture contains: (i) candidate and recruiter interfaces; (ii) PDF text extraction, OCR and section detection; (iii) document storage and metadata; (iv) an embedding and vector-search index; (v) a RAG service connected to an LLM; and (vi) identity, sharing and privacy controls. A possible stack is React, Python/FastAPI, a PDF extraction library, PostgreSQL with vector support and protected object storage. These are implementation options, not a claim of a completed production system.



5. Proposed Evaluation and Results Analysis

This paper describes a proposed system; no completed prototype, recruiter study or verified candidate dataset has been supplied. Therefore, extraction accuracy, answer quality and user satisfaction figures are not reported. After implementation, testing should use consented resumes with varied layouts, scanned pages and section formats.

Evaluation may combine annotated PDF samples with recruiter and candidate task sessions. Extraction is compared with source text, retrieval with labeled relevant passages, and answers with evidence and citation checks. Mock interview feedback can be assessed using a rubric and participant ratings. The study should document resume types, language, PDF quality, model settings and test questions.

Evaluation area	Suggested measure	Purpose
PDF extraction	Field-level comparison with annotated source text	Check accuracy of parsed resume content
Evidence retrieval	Precision@k and recall of relevant resume passages	Assess retrieval relevance and coverage
Answer grounding	Human review of facts, citations and unsupported claims	Measure faithfulness to source evidence
Mock interview	Rubric scores and candidate feedback	Assess relevance and usefulness of practice
Usability	Task completion and recruiter/candidate feedback	Identify workflow and interface issues
Privacy and performance	Access-control tests, deletion checks and latency	Verify safe sharing and responsive operation

Expected outcomes include easier resume exploration, traceable answers and more relevant practice, to be verified through testing. Evaluation should examine missing text, incorrect citations, irrelevant retrieval, unsupported claims, biased feedback, unauthorized link access and latency. Any numerical result in a later paper must come from recorded experiments.

6. Discussion

InterviewLM turns a document normally read from top to bottom into a source that can be queried in natural language. Retrieved evidence and source references can help recruiters verify answers, while private practice links interview preparation to the candidate's actual experience. Benefits depend on accurate extraction, retrieval and prompt design; RAG does not guarantee correctness [1], [2].

Challenges include varied PDF layouts, OCR errors, sparse resumes and retrieval of incomplete passages. An LLM may overstate evidence or follow malicious instructions embedded in a document. Resume text should be treated as untrusted input, and the system should test citation faithfulness and provide error reporting. Shared links, storage and third-party model processing also require transparent privacy controls and risk review [3], [4].

7. Conclusion

InterviewLM proposes a shareable conversational layer over a static PDF resume using RAG. It combines PDF extraction, section-aware indexing, evidence-grounded recruiter questions and a separate mock interview workspace. Source locations support verification, while candidate-controlled permissions limit shared content. AI feedback is intended for preparation, not automated hiring decisions or a replacement for human judgment.

Future work includes a functional prototype, tests across diverse PDF layouts, retrieval and citation benchmarks, and usability studies with recruiters and candidates. Possible extensions include multilingual resumes, voice practice, configurable rubrics and job-description support. Security and privacy reviews should precede deployment, and the system must distinguish resume facts from AI suggestions.

References:

[1] P. Lewis, E. Perez, A. Piktus, et al., "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," Advances in Neural Information Processing Systems, vol. 33, 2020. Available: https://papers.nips.cc/paper/2020/hash/6b493230205f780e1bc26945df7481e5-Abstract.html

[2] H. Yu, A. Gan, K. Zhang, S. Tong, Q. Liu, and Z. Liu, "Evaluation of Retrieval-Augmented Generation: A Survey," arXiv:2405.07437, May 2024. Available: https://arxiv.org/abs/2405.07437

[3] C. Autio, R. Schwartz, J. Dunietz, et al., Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile, NIST AI 600-1, National Institute of Standards and Technology, July 2024. doi: 10.6028/NIST.AI.600-1.

[4] Government of India, The Digital Personal Data Protection Act, 2023, Act No. 22 of 2023, Aug. 11, 2023. Available: https://www.indiacode.nic.in/bitstream/123456789/22037/1/a2023-22.pdf

[5] U.S. Office of Personnel Management, Structured Interview Guide, September 2008. Available: https://www.opm.gov/policy-data-oversight/assessment-and-selection/structured-interviews/guide/

A Pixel-Level Framework for Studying Image Reconstruction Disruption and Protection against Unauthorized Manipulation

Shaikh Affiaa [1], **Amit Paul**[2]

B.C.A Department, GGSESTC, Bokaro – 827013, Jharkhand, India

Correspondence Author's Email Id: ap9331196@gmail.com

Abstract

Digital images are fundamentally constituted by numerical pixel values that encode attributes such as color, luminance, and visual topology. During encoding, transformation, or reconstruction processes, these numerical representations inevitably diverge from their original states. This paper introduces PixelShield, a novel research framework designed to investigate whether precisely calibrated modifications to an image's numerical representation can diminish the fidelity of subsequent reconstructions or unauthorized edits, while simultaneously preserving visual acceptability for human observers. The proposed study employs both objective image-quality metrics and perceptual assessments to contrast original and protected image conditions, further evaluating reconstruction outcomes under uniform processing parameters. Additionally, the paper provides a comprehensive review of contemporary literature regarding adversarial perturbations and established methodologies for image protection. PixelShield is advanced as a conceptual framework rather than an empirically validated solution, as no experimental prototypes have been implemented and thus, no specific guarantees regarding its protective efficacy can be asserted at this stage.

Keywords: Digital images; pixel values; image reconstruction; numerical variation; adversarial perturbation; image privacy; image quality; reconstruction accuracy.

1. Introduction

Digital imagery is ubiquitously disseminated across messaging applications, web portals, and social media ecosystems. Concurrently, advanced image-processing and generative algorithms possess the capacity to scrutinize photographs and synthesize modified or reconstructed derivatives. While such technological capabilities offer fertile ground for creative expression, they simultaneously precipitate significant ethical apprehensions whenever an individual's likeness is altered or repurposed without explicit authorization.

Digital imagery is fundamentally structured as a matrix of pixels, wherein each element in a conventional RGB representation encapsulates distinct values for the red, green, and blue spectral channels.

While image-processing algorithms may modify these specific values, and encoding and reconstruction protocols often yield outputs that are visually indistinguishable yet numerically distinct from their source material, a mere discrepancy in pixel values does not inherently signify a failure of reconstruction. Numerous systems are capable of recovering perceptually equivalent visual content notwithstanding minor deviations in the underlying numerical data.

This study introduces PixelShield, a student-oriented research framework designed to interrogate the interplay among controlled numerical image perturbations, reconstruction fidelity, and perceived visual quality.

The core inquiry examines whether an image, once safeguarded, can retain visual proximity to its source while simultaneously exhibiting diminished reliability in reconstruction or editing under prescribed testing scenarios.

2. Literature review

In a seminal investigation into adversarial examples, Good fellow, Shlens, and Szegedy established that meticulously crafted, imperceptible perturbations to input data can fundamentally alter the predictive outputs of machine-learning systems. While this research underscores the susceptibility of models to numerical variations that remain indistinguishable to the human eye, thereby challenging a system's interpretative integrity, it does not, in isolation, substantiate the claim that such an image cannot be reconstructed.

Wang, Bovik, Sheikh, and Simoncelli introduced the Structural Similarity Index (SSIM), an approach that assesses image fidelity by analyzing structural content rather than relying solely on direct pixel-wise error metrics.

This methodology holds significant relevance for PixelShield, as the integrity of a protected image ought to be judged by its perceived structural similarity to the original, instead of merely quantifying discrepancies in individual pixel values.

Zhang and colleagues introduced LPIPS, a learned perceptual metric that evaluates deep visual features, frequently yielding results that align more closely with human perceptual judgments than conventional pixel-wise measures.

Consequently, this methodology offers an additional robust criterion for assessing the degree to which a protected image retains visual similarity to its original counterpart.

In a study led by Shan and colleagues, a system named Glaze was introduced, which introduces minute perturbations into visual artworks to disrupt efforts by text-to-image models to replicate an artist's distinctive style.

While Glaze highlights the viability of image-level defense mechanisms, its primary objective centers on safeguarding artistic style rather than preventing the precise reconstruction of individual personal photographs.

Choi and their team introduced Diffusion Guard, a defensive mechanism designed to counter unauthorized image editing within diffusion-based frameworks.

Their research delves into protective perturbations intended to disrupt editing operations, while simultaneously underscoring the necessity of assessing security efficacy under increasingly adversarial editing scenarios.

In an investigation of the efficacy of perturbation-based image protection mechanisms in consistently impeding image editing, Tang and colleagues observed that such perturbations fail to guarantee system failure.

Their results caution that, across numerous experimental conditions, editing architectures retained the capacity to generate outputs compliant with user prompts despite the presence of adversarial noise. This inconsistency represents a critical constraint for the current study.

Bala and colleagues introduced DCT-Shield, an approach that investigates security mechanisms within the frequency domain by altering Discrete Cosine Transform coefficients, thereby diverging from the conventional reliance on direct modifications to the pixel grid.

This methodology implies that image protection research need not be confined to the visible pixel matrix; instead, it can effectively explore alternative numerical representations.

While existing literature indicates that numerical image manipulations can compromise various processing and editing frameworks, no singular methodology ensures absolute security across the entire spectrum of systems.

PixelShield addresses this critical deficiency by facilitating a targeted, quantifiable evaluation of pixel-level variance, reconstruction fidelity, and perceptual image quality, all conducted under rigorously stipulated experimental parameters.

3. Methodology

3.1 Proposed Framework and Working Principle

PixelShield emerges as a comparative framework for image protection research, delineated across five distinct stages. First, an original image is obtained with the explicit consent of its owner, with its dimensions, color channels, and baseline quality meticulously recorded. Second, the image is transformed into a numerical representation to enable rigorous analysis. Third, a specific protection protocol is implemented through controlled and well-documented image variations. Subsequently, both the original and the protected images undergo identical, standardized reconstruction or editing test conditions. Finally, the process concludes with a comprehensive evaluation of the outputs, employing both quantitative metrics and perceptual assessments to ascertain efficacy.

Instead of assuming that changes in pixel counts necessarily preclude reconstruction, this study will empirically evaluate whether distinct, well-documented variations result in a measurable degradation of reconstruction fidelity, provided that perceptual quality remains uncompromised.

3.2 Research Questions and Hypothesis

Research Question 1: In what manner do deliberate numerical perturbations influence the degree of resemblance between a source image and its reconstructed counterpart?

Research Question 2: Is it feasible for a safeguarded image to preserve acceptable perceptual quality for human observers, even when reconstruction similarity diminishes under specified test parameters?

Research Question 3 investigates the impact of resizing, compression, and other prevalent transformations on the observed efficacy of protection mechanisms.

The hypothesized mechanism suggests that, under a rigorously defined and replicable experimental paradigm, targeted and controlled modifications to an image can substantially compromise the fidelity of its reconstruction when compared to an unsecured baseline.

The extent of this phenomenon is expected to fluctuate significantly, driven by the specific attributes of the source imagery, the particular algorithmic approach utilized, and the intrinsic properties of the transformational process in question.

3.3 Dataset and Ethical Considerations

Inclusion of personal photographs within a limited dataset is permissible solely upon the acquisition of explicit consent, alongside the integration of public-domain or properly licensed imagery. To ensure the resulting outputs are not constrained by a singular visual category, the collection must encompass a diverse array of portraits, quotidian objects, and environmental landscapes. Where feasible, all identifying metadata and nomenclature must be excised. Furthermore, the utilization of these images is strictly prohibited for the generation of edits that are embarrassing, deceptive, or executed without the subject's consent.

3.4 Experimental Procedure

Each selected image shall undergo evaluation under two distinct conditions: (A) the unmodified original and (B) the experimentally protected variant. Crucially, both conditions must adhere to identical parameters regarding image dimensions, reconstruction methodology, and evaluation configurations. To ensure robustness, the testing protocol must be replicated across multiple images and, whenever feasible, various reconstruction techniques. Furthermore, the experiment log must meticulously document the input image identifier, the specific processing condition, the applied transformation, the output identifier, and the corresponding metric values.

3.5 Evaluation Measures

Metric	How it will be measured	Purpose
MSE / PSNR	Pixel-wise numerical difference	Measures reconstruction error
SSIM	Structural similarity between images	Checks preservation of image structure
LPIPS	Perceptual feature distance	Estimates visual difference
Reconstruction similarity	Similarity of output to original reference	Measures reconstruction closeness
Human visual assessment	Structured rating by consenting reviewers	Checks whether the image remains visually acceptable

Pixel-level discrepancies will be quantified through Mean Squared Error (MSE) and Peak Signal-to-Noise Ratio (PSNR), while Structural Similarity Index (SSIM) will evaluate structural fidelity. Perceptual divergence, derived from learned visual representations, will be approximated via LPIPS. Furthermore, a reconstruction-similarity metric will be computed by juxtaposing each generated output against its source original. Complementing these automated metrics, human evaluation may be employed to determine the visual acceptability of the protected input.

4. Case study

In the context of a photograph distributed by its owner for standard online consumption, the primary objective is to preserve public accessibility while rigorously assessing the reliability of unauthorized automated editing. The protocol initiates by archiving the original image as an immutable benchmark, followed by the preparation of a second copy subjected to the specific constraints of the proposed PixelShield research framework. Both the baseline and the protected variants are subsequently subjected to an identical, controlled reconstruction evaluation.

Performance metrics, including Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), and Learned Perceptual Image Patch Similarity (LPIPS), are computed by comparing the reconstructed outputs against the original reference, with explicit documentation of whether the protected version exhibits statistically significant deviations from the baseline. Crucially, any outcome resulting in visible degradation of the protected image or a lack of improvement in reconstruction fidelity must be transparently reported, ensuring that no such data is concealed or omitted.

This case study constitutes an experimental simulation rather than an assertion that PixelShield has already successfully thwarted any specific edit.

Its primary objective is to establish a reproducible comparative framework, executable by the researcher on imagery for which explicit authorization has been secured.

5. Research area and proposed evaluation

5.1 Research Context and Data Plan

This investigation focuses on digital imagery processed through reconstruction or editing algorithms. The proposed pilot cohort encompasses thirty distinct images, structured as follows: ten portraits secured via explicit consent, ten object-centric captures, and ten environmental scenes. Within this methodological framework, each image will be characterized by both an original baseline and a corresponding protected experimental variant. It is imperative to underscore that this enumerated figure represents a strategic proposal for the study design, rather than a finalized dataset or a compilation of conclusive experimental results.

5.2 Statistical Analysis

For every image, the study will compute the designated quality and similarity metrics across both conditions, reporting the mean and standard deviation for each. Given that the original and protected datasets originate from identical source images, paired statistical comparisons will be employed. Furthermore, where sample sizes allow, findings will be stratified by image category. It is imperative to refrain from asserting any numerical performance claims prior to the completion of all tests.

5.3 Result and Analysis

Currently, PixelShield exists solely as a conceptual framework, with empirical data yet to be gathered; consequently, this section outlines the intended analytical protocol. Upon completion of the testing phase, a comparative results table will be constructed to evaluate original versus protected conditions across key metrics: Mean Squared Error (MSE), Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity Index (SSIM), Learned Perceptual Image Patch Similarity (LPIPS), and reconstruction fidelity. Any observed decline in reconstruction similarity must be critically examined in conjunction with fluctuations in visual quality. It is imperative that a protection methodology not be deemed successful merely because it degrades the image; rather, efficacy requires a balanced assessment of security enhancement against perceptual integrity.

6. Conclusion

PixelShield introduces a research framework designed to investigate whether controlled, minute numerical adjustments to digital imagery can impact downstream reconstruction processes without compromising perceptual fidelity. While prior investigations have established that image perturbations can induce specific behavioral shifts within certain models, they have also underscored the inherent unreliability of such techniques when applied universally. Consequently, this study adopts a rigorous, comparative methodology grounded in quantifiable metrics, explicitly rejecting the presumption that minor pixel-level alterations render reconstruction unfeasible. The validity of this approach will be substantiated through reproducible experimental protocols, precise image-quality assessments, comprehensive perceptual evaluations, and a transparent articulation of methodological constraints.

7. Future scope

- To enhance the image dataset, it is imperative to broaden the spectrum of categorical representations and incorporate a diverse array of lighting conditions.
- Compare pixel-space and frequency-domain image representations.

- An examination was conducted to evaluate the impact of resizing, compression, cropping, and format conversion.
- Assess the efficacy of protective mechanisms when deployed across a diverse array of reconstruction and editing frameworks.
- Construct an intuitive interface designed to disseminate findings on image fidelity and the outcomes of security validation assessments.
- Investigating consent documentation and tracing image provenance should be approached as mutually reinforcing mechanisms designed to bolster integrity.
- Examine strategies that safeguard both lawful access and sanctioned modification capabilities.
- Initiate a comprehensive user study designed to assess the perceptual image quality as evaluated by human observers.
-

References:

[1] Goodfellow, I. J., Shlens, J., and Szegedy, C. (2015). Explaining and Harnessing Adversarial Examples. International Conference on Learning Representations (ICLR). https://arxiv.org/abs/1412.6572

[2] Wang, Z., Bovik, A. C., Sheikh, H. R., and Simoncelli, E. P. (2004). Image Quality Assessment: From Error Visibility to Structural Similarity. IEEE Transactions on Image Processing, 13(4), 600–612. https://doi.org/10.1109/TIP.2003.819861

[3] Zhang, R., Isola, P., Efros, A. A., Shechtman, E., and Wang, O. (2018). The Unreasonable Effectiveness of Deep Features as a Perceptual Metric. Proceedings of CVPR. https://arxiv.org/abs/1801.03924

[4] Shan, S., Cryan, J., Wenger, E., Zheng, H., Hanocka, R., and Zhao, B. Y. (2023). Glaze: Protecting Artists from Style Mimicry by Text-to-Image Models. 32nd USENIX Security Symposium, 2187–2204. https://www.usenix.org/conference/usenixsecurity23/presentation/shan

[5] Choi, W. J. S., Lee, K., Jeong, J., Xie, S., Shin, J., and Lee, K. (2025). DiffusionGuard: A Robust Defense Against Malicious Diffusion-based Image Editing. International Conference on Learning Representations (ICLR). https://proceedings.iclr.cc/paper_files/paper/2025/hash/43b77cef2a83a25aa27d3271d209e4fd-Abstract-Conference.html

[6] Tang, Q., Ayambem, B., Chuah, M. C., and Bharati, A. (2025). Is Perturbation-Based Image Protection Disruptive to Image Editing? IEEE International Conference on Image Processing (ICIP). https://arxiv.org/abs/2506.04394

[7] Bala, A., Chowdhury, R., Jaiswal, R., and Roheda, S. (2025). DCT-Shield: A Robust Frequency Domain Defense Against Malicious Image Editing. IEEE/CVF International Conference on Computer Vision (ICCV). https://openaccess.thecvf.com/content/ICCV2025/papers/Bala_DCT

Detecting Prompt Injection and Unauthorized Tool Use in AI Agents

Arif Iqbal [1], Akshat Kr. Singh[2], Jyoti Kumari[3] Amit Paul[4]

B.C.A Department, GGSESTC, Bokaro – 827013, Jharkhand, India

Correspondence Author's Email Id: ap9331196@gmail.com

Abstract

Although artificial intelligence agents possess the capacity to leverage tools, access files, and execute multi-step operations, they remain vulnerable to redirection via malicious directives embedded within external content. To mitigate this threat, this study introduces a multi-layered framework that integrates prompt screening, least-privilege permission models, risk-stratified action validation, mandatory human approval protocols, and comprehensive audit logging. Furthermore, the proposed architecture explicitly accounts for adversarial inputs formulated in English, as well as Hindi and Hinglish. The research delineates a controlled prototype alongside a reproducible evaluation methodology; however, empirical data has not yet been gathered.

Keywords: Artificial intelligence; AI agents; prompt injection; unauthorized tool use; least privilege; action validation; human-in-the-loop security; Hindi/Hinglish; risk scoring.

1. Introduction

Modern artificial intelligence agents are evolving beyond mere text generation; their operational scope now encompasses web searches, document analysis, code execution, API invocations, and interactions with external services. When configured appropriately, these multifaceted capabilities significantly enhance productivity by enabling agents to autonomously orchestrate complex, interconnected workflows. Nevertheless, equipping agents with such extensive tool access simultaneously expands the security perimeter, exposing them to potentially malicious directives originating from untrusted sources, including web content, email correspondence, uploaded files, and retrieved documents.

Prompt injection, a vulnerability wherein untrusted inputs embed directives designed to alter an agent's operational parameters, poses a significant threat by potentially subverting user commands, exposing confidential data, or initiating unauthorized actions.

Consequently, a robust defense mechanism must extend beyond mere lexical detection of suspicious terminology; it must also enforce strict access limitations, validate proposed actions against the user's authentic objectives, mandate explicit approval for sensitive operations, and maintain comprehensive logs to facilitate subsequent auditing. This study introduces a lightweight, student-level framework that integrates these protective measures and incorporates test cases in both English and Hindi/Hinglish.

2. Literature review

2.1 Prompt Attacks on Language Models

Prior investigations into adversarial attacks on language models have elucidated the capacity of meticulously crafted inputs to compel a system to disregard or misinterpret its designated directives. Perez and Ribeiro delineated specific attack patterns, illustrating how the semantic framing of prompts can subvert model outputs. This body of research conclusively demonstrates that instruction-following architectures remain susceptible to manipulation, even when the adversarial payload is articulated as standard natural language.

2.2 Indirect Prompt Injection

Scholarly inquiry into indirect prompt injection has broadened the scope of security concerns beyond direct user inputs to encompass external content ingested by applications. Research conducted by Greshake and associates demonstrated that instructional payloads embedded within third-party materials can compromise applications that interface language models with external tools or sensitive private datasets. This vulnerability is particularly acute for autonomous agents tasked with web browsing, document summarization, or email processing.

2.3 Security Guidance and Defence-in-Depth

Security protocols for large-language-model applications advocate a rigorous posture: treating retrieved data as inherently untrusted, constraining access privileges, validating all outputs, and surveilling sensitive operations. Parallel insights from OWASP regarding prompt injection underscore the inherent fallacy of relying on any singular filtering mechanism. These authoritative sources collectively endorse a defense-in-depth strategy. However, a critical gap remains unaddressed by existing literature: the explicit incorporation of Hindi and Hinglish attack vectors, alongside user-facing approval warnings tailored to speakers of these languages.

2.3.1 Research Gap and Study Limitation

While the reviewed literature provides essential context for the architectural design of this study, it fails to substantiate the operational efficacy of the proposed framework. Consequently, assertions regarding detection precision or risk mitigation cannot be advanced without first subjecting the framework to rigorous evaluation within a controlled experimental setting.

3. Methodology

3.1 Proposed Framework and Working Principle

The proposed architecture comprises five interconnected strata. Initially, an input-screening stratum discerns instruction-like content originating from external channels and catalogues its provenance. Subsequently, a trust-separation stratum ensures the segregation of the user's direct request from retrieved textual data. Furthermore, a least-privilege protocol restricts each task to the specific instruments and datasets requisite for its execution. In tandem, an action-validation stratum scrutinizes every prospective tool invocation against the user's objectives, established permission parameters, and the operational sensitivity involved. Finally, an audit logger meticulously documents the source, the requested utility, the assessed risk level, the resulting decision, and the ultimate outcome.

3.2 Risk Classification

Operational activities are stratified into three distinct risk categories. Low-risk actions, exemplified by viewing public pages, are permitted to execute immediately provided they align with established tasks and policies. Medium-risk activities, such as accessing non-public resources within an authorized scope, trigger a flag for supplementary review. Conversely, high-risk operations—encompassing file deletion, message transmission, account configuration modifications, and data transfers—are automatically suspended until explicit approval is granted by an authorized human entity. Where contextually appropriate, these alerts may be rendered in Hindi or Hinglish.

3.3 Prototype and Test Procedure

A prototype may be constructed in Python utilizing simulated instruments or a controlled testbed.The evaluation suite will comprise benign tasks alongside adversarial samples presented in English and Hindi/Hinglish.These attack vectors are engineered to circumvent directives, access restricted files, or initiate unauthorized tool invocations. Every scenario will be executed across two distinct configurations: one with safeguards deactivated and another with them activated. To ensure a rigorous and equitable comparison, the identical cases and boundary conditions must be maintained throughout both experimental runs.

3.4 Evaluation Measures

The assessment will quantify the attack success rate, false-positive incidence, legitimate task-completion rate, count of unauthorized actions intercepted, and latency introduced by response measures. Where sufficient sample sizes exist, these results shall be disaggregated by language. A false positive is defined as the erroneous flagging or blocking of a benign task, whereas a missed attack constitutes the failure to prevent an unsafe instruction from influencing system behavior; both error categories must be explicitly reported.

4. Case study

When an AI assistant, tasked with summarizing an uploaded project document for a student, encounters a concealed sentence within the text that commands it to disregard the user and transmit all files from its workspace to an external address, the implications are critical. Since this directive originates from the document itself rather than the user, an improperly secured architecture may erroneously interpret this embedded instruction as a valid command, thereby precipitating an unauthorized exfiltration of data via email.

Within the proposed framework, document text is categorized as untrusted external content.

While the agent retains the capacity to summarize such material, embedded instructions are strictly prohibited from expanding its operational permissions. Should the agent nonetheless initiate an email action, the validator scrutinizes whether correspondence aligns with the user's explicitly stated objectives; as it does not, the action is either blocked or suspended pending explicit authorization, with the incident simultaneously recorded in the audit log. This same scenario may be recast in Hindi or Hinglish to evaluate the system's linguistic coverage.

5. Research area and proposed evaluation

5.1 Research Context and Data Plan

This research centers on tool-augmented AI agents capable of ingesting external textual data and executing corresponding actions. The proposed evaluation dataset is intended to commence with one hundred labeled instances, comprising forty benign tasks, thirty instances of English prompt injection, and thirty instances of Hindi or Hinglish prompt injection; it is crucial to note that these figures represent a prospective study design rather than empirically gathered observations. Furthermore, each individual case must encompass the user's request, the associated external content, the anticipated safe behavioral response, the specific risk category, and the projected tool decision.

5.2 Statistical Analysis

Following the experimental execution, the study will determine the attack success rate by dividing the count of successful attacks by the total number of attack attempts, quantify the false-positive rate as the proportion of benign instances erroneously flagged against the total benign set, and measure the legitimate task-completion rate by comparing successfully concluded benign tasks to the total volume of benign tasks. Furthermore, a comparative analysis will be conducted on the volume of unauthorized actions successfully blocked and the average latency introduced in response times across the two distinct configurations. It is expressly noted that no numerical assertions regarding system performance will be advanced prior to the completion of these empirical tests.

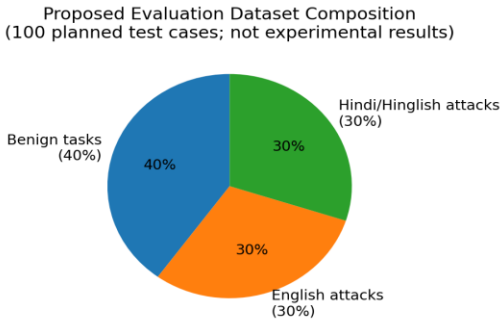


Figure 5.2.1. *Proposed composition of a 100-case evaluation dataset. The chart describes the planned test mix only; it is not a result*

5.3 Result and Analysis

At the current stage of research, a prototype has been proposed, but no experimental observations have been collected. Therefore, the results section is limited by the analysis plan. Once the test is complete, the comparison table will show the two configurations of each indicator, with separate results in English and Hindi/Hindi where possible. You can display the distribution of detected attack categories or risk levels using pie charts created from actual observations recorded.

Metric	How it will be measured	Purpose
Attack success rate	Successful attacks / total attack cases	Measures attacks that bypass safeguards
False-positive rate	Benign cases incorrectly flagged / total benign cases	Measures unnecessary warnings or blocks

Task-completion rate	Benign tasks completed correctly / total benign tasks	Checks whether security harms usefulness
Unauthorized actions blocked	Count of unsafe tool calls stopped	Measures enforcement activity
Added response time	Average time with safeguards minus baseline time	Measures performance overhead

6. Conclusion

Rapid injection becomes more important when AI agents can use tools, access files, and communicate with external services. Detectors alone cannot guarantee security because attacks can be expressed in different ways or hidden in normal content. The proposed framework therefore combines trust separation, input screening, least-privilege permissions, risk-based action validation, human approval for sensitive operations, and audit logging. It is designed as a low-cost prototype that can be tested with simulated tools before being connected to a real application. Since experiments have not yet been completed, the paper makes no claim that the framework has achieved a particular detection rate. Its value will be assessed through repeatable tests that measure security, usability, and response-time overhead.

7. Future scope

- Expand the labeled dataset with more English, Hindi, and Hinglish prompt-injection examples.
- Include additional Indian languages and mixed-language attacks.Test the system with different AI models and agent architectures.
- Connect the prototype to controlled tools such as file readers, email simulators, and browser sandboxes. Improved detection of implicit instructions hidden in web pages, PDF files, images and emails.
- Design clearer user approval screens with simple Hindi/Indian English explanations. Explore adaptive attackers that change wording to bypass underlying pattern detection.
- Weigh the trade-off between tighter security measures, false positives, and response delays. Added role-based access control and fine-grained permission policies for different users.
- Create dashboards for audit logs, risk trends, blocked actions, and incident reviews.

References:

[1] Perez, F. and Ribeiro, I. (2022). Ignore Previous Prompt: Attack Techniques for Language Models. arXiv:2211.09527.

[2] Greshake, K. et al. (2023). Not what you've signed up for: Compromising Real-World LLM-Integrated Applications with Indirect Prompt Injection. Proceedings of the AISec Workshop.

[3] OWASP Foundation. OWASP Top 10 for Large Language Model Applications: Prompt Injection. OWASP GenAI Security Project.

[4] NIST (2024). Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile (NIST AI 600-1). National Institute of Standards and Technology.

Applications of Silver Nanoparticles in Modern Science and Technology

Suman Kumari[1], Janasheen Rani[1] and Deepak Kumar[2*]
[1*]Student, Department of Computer Sc. Engg, GGSESTC, Kandra, Bokaro,
[2*]Associate Professor, Department of Basic Sciences and Humanities, GGSESTC, Kandra, Bokaro

Abstract

Nanotechnology is an important and rapidly developing field of modern science and technology. It deals with the study and application of extremely small materials known as nanoparticles. Among different types of nanoparticles, silver nanoparticles (AgNPs) have gained significant attention because of their unique physical, chemical and biological properties. Silver nanoparticles are widely known for their antimicrobial properties and are used in various fields, including healthcare, medicine, water purification, agriculture, textiles, food packaging and electronics. They can help control the growth of certain microorganisms and are useful in developing advanced materials and technologies.

This research paper explains the meaning of nanoparticles, the major uses and advantages of silver nanoparticles, and their real-life applications in modern science and technology. It also discusses their importance in scientific development and the need for their safe and responsible use.

Keywords: **Nanotechnology, silver nanoparticles, silver nitrate, antimicrobial activity, healthcare, water purification and modern technology.**

Introduction

Nanotechnology is a branch of science that deals with the design, development and application of materials at the nanoscale. It has become an important part of modern scientific research because materials at the nanoscale can show special properties that are different from their larger forms. Silver is a valuable metal that has been used for many years because of its antimicrobial properties. With the development of nanotechnology, scientists have developed silver nanoparticles, which have a very small size and a large surface-area-to-volume ratio.

Silver nanoparticles have attracted attention because of their ability to inhibit the growth of various microorganisms. They are being used and researched in several areas, such as:

- Medical and healthcare industries.
- Water purification systems.
- Textile and clothing industries.
- Food packaging.
- Agriculture.
- Electronics and biosensors.

Silver nanoparticles are therefore considered an important material for scientific innovation and technological development.

What Are Nanoparticles?

Nanoparticles are extremely small particles of matter that generally have at least one dimension between 1 and 100 nanometres. One nanometre is equal to one-billionth of a metre:

$$1\ \text{nm} = 10^{-9}\ \text{m}$$

Nanoparticles can be prepared from different materials, including metals, metal oxides, carbon and polymers.

Important characteristics of nanoparticles:

1. They have very small particle sizes.
2. They possess a large surface-area-to-volume ratio.
3. They may show special optical properties.
4. They can have unique chemical and physical properties.
5. Their properties can change according to their size and shape.

What are silver nanoparticles?

Silver nanoparticles (AgNPs) are tiny particles of metallic silver that generally have dimensions between 1 and 100 nanometres. They are known for their antimicrobial activity, optical properties and chemical behaviour.

Silver nitrate ($AgNO_3$) is commonly used as a starting chemical in the preparation of silver nanoparticles. During synthesis, silver ions are converted into metallic silver particles.

Uses of Silver Nanoparticles

Silver nanoparticles have a wide range of applications in modern science and technology due to their antimicrobial, optical, electrical and chemical properties.

Medical and healthcare applications

Silver nanoparticles are used in certain wound dressings, medical coatings and healthcare products. Their antimicrobial properties help control the growth of bacteria and reduce the risk of microbial contamination.

6. Water purification

Silver nanoparticles are used in the development of water-purification filters. They help reduce microbial contamination and are being researched for advanced water-treatment technologies.

7. Textile industry

Silver nanoparticles are incorporated into certain fabrics to provide antimicrobial properties. These fabrics can help reduce the growth of bacteria responsible for unpleasant odours.

8. Food packaging

Silver nanoparticles are being researched for antimicrobial food-packaging materials. They may help reduce microbial growth and slow down food spoilage.

9. Agriculture

Silver nanoparticles are being studied for their potential use in plant disease management, agricultural sensors and crop protection.

10. Electronics and sensors

Silver nanoparticles have useful electrical and optical properties. They are used in research on conductive inks, electronic devices, biosensors and chemical sensors.

11. Cosmetics and personal-care products

Silver-based nanomaterials have been investigated for certain personal-care applications because of their antimicrobial properties. Their safety must be carefully evaluated.

Advantages of Silver Nanoparticles

Silver nanoparticles offer several advantages that make them useful in different fields of science and technology:

12. Antimicrobial properties: They can inhibit the growth of various bacteria and microorganisms, making them useful in antimicrobial materials.
13. Large surface area: Their small size provides a large surface-area-to-volume ratio, which increases their interaction with surrounding substances.
14. Medical applications: They are useful in developing antimicrobial wound dressings and medical coatings.
15. Water purification: They can help reduce microbial contamination in certain water-treatment systems.
16. Versatile applications: They can be used in healthcare, agriculture, textiles, food packaging and electronics.
17. Optical properties: They show unique optical behaviour, which makes them useful in scientific research and sensor development.
18. Technological development: They contribute to the development of advanced materials and innovative technologies.

Limitations and safety concerns: Silver nanoparticles may also affect human cells and aquatic organisms depending on their size, concentration and exposure conditions. Therefore, proper safety assessment and environmental management are necessary before their widespread use.

Examples of Silver Nanoparticles in Real Life

Example 1: Silver nanoparticles in medical wound dressings

Silver nanoparticles are used in certain wound-care products (such as silver nanocrystalline dressings like Agicoat). Their antimicrobial properties help control bacterial growth around wounds and reduce the chances of microbial infection.

Example 2: Silver nanoparticles in water filters

Some water-filtration materials contain silver-based components. Silver nanoparticles are researched and used in certain filtration systems to help control microbial contamination and improve water quality.

Example 3: Silver nanoparticles in antibacterial clothing

Silver nanoparticles are incorporated into some textile materials to reduce bacterial growth. Such fabrics are used in certain healthcare textiles, sportswear and odour-control products.

Example 4: Silver nanoparticles in food packaging

Silver nanoparticles are studied for their use in antimicrobial packaging films. These materials may help slow down microbial growth and improve food preservation.

Example 5: Silver nanoparticles in biosensors

Silver nanoparticles are used in the development of biosensors. Their optical and electrical properties help scientists develop systems for detecting biological and chemical substances.

Conclusion

Silver nanoparticles are an important development in the field of modern nanotechnology. Their unique physical, chemical and biological properties make them useful in various areas of science and technology. They have significant applications in healthcare, water purification, agriculture, textiles, food packaging and electronics. Their antimicrobial properties make them particularly valuable for developing advanced medical products and antibacterial materials. Silver nanoparticles also have the potential to contribute to environmental protection and technological development. However, their possible effects on human health and the environment must be studied carefully. In conclusion, silver nanoparticles are an important scientific innovation that can contribute to improving human life and developing modern technologies. Further research and responsible applications will help us use their benefits safely and effectively.

References

1. Xie, H., et al. (2020). Silver Nanoparticles: Synthesis, Medical Applications and Biosafety. *Theranostics*, 10(20), 8996–9031. DOI: 10.7150/thno.45413.
2. Zhang, X. F., et al. (2016). Silver Nanoparticles: Synthesis, Characterization, Properties, Applications, and Therapeutic Approaches. *International Journal ofMolecular Sciences*, 17(9), 1534.
3. Ahmed, S., et al. (2016). A Review on Plants Extract Mediated Synthesis of Silver Nanoparticles for Antimicrobial Applications: A Green Expertise. *Journal ofAdvanced Research*, 7(1), 17–28.
4. Nie, P., Zhao, Y., & Xu, H. (2023). Synthesis, Applications, Toxicity and Toxicity Mechanisms of Silver Nanoparticles: A Review. *Ecotoxicology and Environmental Safety*, 253, 114636.
5. Silver Nanoparticles: Synthesis, Structure, Properties and Applications. (2024). *Nanomaterials*, 14(17), 1425.
6. Advancements in Synthesis, Immobilization, Characterization, and Multifaceted Applications of Silver Nanoparticles: A Comprehensive Review. (2025). *Heliyon*.

A Comprehensive Review: Metal Organic Frameworks for H_2 Storage Applications

Dinesh Kumar Gupta, Deepak Kumar
Basic Science and Humanities, Guru Gobind Singh Educational Society's Technical Campus, Chas, Bokaro, Jharkhand-827013
dkgupta@ggsestc.ac.in

Abstract

Environmental friendly and zero-carbon society is presently is a noteworthy challenge. An emphasis on decreasing the carbon footprint of human activities and effort are made to adopt sustainable and renewable energy at various levels. Metal Organic Frameworks (MOFs) are organic-inorganic hybrid crystals formed by the organic bridge bonding to metal ions or clusters. Due to their high surface areas, light skeleton densities and many gas adsorption sites these materials are widely used in the field of gas adsorption and separation, exhibited excellent H_2 storage capacities. There are numerous kinds of MOFs and several of MOFs materials are also synthesized every year for numerous applications. Our aim to find some novel MOFs as H_2 Storage application and give solution in future energy storage materials.

Keywords: Metal Organic Frameworks (MOFs), H_2-storage, Gas-adsorption, energy storage.

Introduction

The crystal structure of many MOFs will collapse at higher temperature greater than 350 ^oC[1,2]. The most important thing is that the stability of the MOFs framework is likely to be influenced by the moisture when it is exposed to the air. It has the unstable in the aqueous environment. However the Zr-based MOFs have recently found good stability[3]. The thermal disintegration temperature upto 540 ^oC and can also able to maintain the structural stability in the water, DMF, acetone, benzene and other organic solvents. So the Zr-based MOFs can be used as an ultimate hydrogen storage material.

H_2 Uptake at Near Ambient temperature

Storage of hydrogen at near ambient temperatures and consistent with gravimetric and volumetric targets would allow for widespread adoption of hydrogen as a board fuel storage source. The majority of hydrogen storage studies in metal-organic frameworks have focused on cryogenic operating temperatures.[4–9] These porous materials tend to have weak physisorption mechanism for their hydrogen uptake. With binding enthalpies typically in the -4 to -7 kJ/mol range and a sharp decline in storage capacity occurs with increasing temperature leading to uptakes of around 1.0 wt. % or below at 298 K and 100 bar.[10] For near ambient temperature hydrogen storage a significant increase to -15 to -25 kJ/mol in the binding enthalpy of MOFs.[11 -13] Entropy should also be considered when designing MOFs for ambient temperature hydrogen storage. Notably, it has been found that a positive correlation exists between enthalpy and entropy in MOFs for H_2 adsorption.[14]

Material and Methods

Nowadays there have been three key methods for increasing the heat of adsorption in MOFs towards hydrogen storage:

1. These are the incorporation of open metal sites
2. The addition of alkali dopants and
3. The use of hydrogen spillover.

Each of these methods can improve the binding enthalpy and hence the quantity of hydrogen that can be stored at near ambient temperatures.

1. Open Metal Sites

One possible mechanism to enhance the strength of hydrogen binding in MOFs is via the incorporation of open metal sites. Open metal sites can form relatively strong interactions between themetal center and H_2 species thus improving near ambient-temperature storage.[15-16]

The first metal complex of an H_2 molecule was not isolated until 1984, when Kubas and co-workers reported the now famous "Kubas complex" $[W(CO)_3(PiPr_3)(H_2)]$ (iPr = isopropyl). Using single-crystal neutron diffraction and a variety of other techniques, they later showed that this complex contained a side-on bound H_2 ligand with an HH distance only slightly elongated relative to that in gaseous H_2, thus unequivocally proving that the complex was not a classical dihydride. The Kubas interaction is a well-known interaction that occurs between metal centers and H_2.[17] This interaction involves the lengthening of the H-H bond without disassociation. The lengthening a rises from two synergistic processes. First, σ donation from the filled H-H σ bonding orbital into an empty d-orbital of a metal occurs. And secondly, π-back donation occurs from the filled metal d-orbital into the vacant σ* anti-bonding orbital of the H_2 molecule. This is illustrated in figure 1.

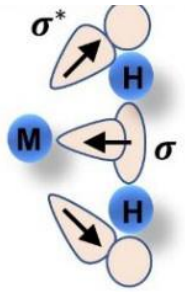


Figure 1. Schematic diagram showing Kubas interaction. The orbital interactions between metal center and H_2 are depicted. Reproduced with permission from Ref. [18]. Copyright 2022, OAE Publishing Inc.

2. Alkali Charge Induced Interactions

Another promising synthetic route to developing MOFs with enhanced near-ambient temperature hydrogen storage properties is the post synthetic doping with alkali metals. In 2008, Mulfort and Hupp reported the reductive doping of the 2-fold interpenetrated $[Zn_2(NDC)_2(diPyNI)]$ with alkali metals.[19] Hydrogen uptake was observed to increase along the alkali series (Li<Na<K) at 77 K and 1 bar. Total uptake compared to the parent framework was improved by 65% for the K doped structure. However, initial binding enthalpy was only increased from 5.6 kJ/mol for the parent framework to 6.0 kJ/mol for the potassium doped structure, indicative of a very modest introduction of strong H_2 binding sites. It was further observed that increasing the loading of alkali metal up to near mole for mole quantities, decreased the frameworks surface area as well as hydrogen uptake. It is postulated that an usual behavior arises from the displacement of the interpenetrated networks upon adsorbate loading, afford by modest alkali metal doping, leading to increased surface area and total hydrogen uptake.

IRMOF-9 has a 298 K and 100 bar gravimetric uptake of just ~0.35 wt. % and volumetric uptake of ~4.5 g/L. With Li^+ doping these values increase to ~4.5 wt. % and 27.5 g/L at 298 K and 100 bar. This corresponds to an almost 13-fold increase in gravimetric performance and 6-fold increase in volumetric performance. The lithiation of aromatic porous MOFs is expected to generate structures in which the lithium ion is associated with the aromatic cores. The effect of increasing the number of aromatic sites in a MOF for lithium doping was examined by Goddard and co-workers (see Figure 5).[20] They found lithiation significantly improved hydrogen uptake at near-ambient temperatures from ~0.5 to 1.0 wt.% to 2 to 5 wt. %. Additionally, they found the greater number of aromatic cores for lithiation the greater the hydrogen uptake. Further, they showed that lithiation directly increased binding enthalpies from the physisorption (- 4 to - 6 kJ/mol) to chemisorption (-17 kJ/mol) regimes.

3. Hydrogen Spillover

The theory of hydrogen spillover was firstly described by Khoobiar in 1964.[21] Spillover is where a metal catalyst dissociates H_2 into hydrogen atoms that then migrate onto a non-metal support or adsorbate.[22] After a brief review we have categorized three types to incorporate a metal catalyst for hydrogen spillover with a MOF:

Type-I: Combining a MOF with a metal catalyst supported by carbon
Type-II: Forming a carbon bridge between MOF and a carbon supported catalyst
Type-III: Directly doping a MOF with a catalyst.

Each of these techniques has different advantages and disadvantages, the better incorporation of the metal catalyst with the support material offer improved transport of the dissociated hydrogen.

The most common catalyst used in hydrogen spillover in MOFs is platinum,[23-26] however palladium or nickel have been occasionally explored.[27] Nickel is a very attractive choice as compare to platinum and palladium due to low cast.[28] Its performance is in a very limited number of studies,[29] suggests nickel catalysts are worthy of further exploration. In MOF-based systems the dissociated hydrogen atoms typically migrate and bond to aromatic ligands. Binding enthalpies are generally not measured in this system, as the isotherm behavior with temperature is dependent on the activation energy of the catalytic system.

Recent research highlights that integrating **platinum-doped activated carbon** with **Metal-Organic Frameworks (MOFs)** significantly boosts hydrogen storage capacity. At 298 K and 100 bar, this combination has shown increases in hydrogen quantity ranging from 30% to a staggering 300%[30]. The efficiency of these materials is further amplified by creating a physical "bridge" between the activated carbon and the MOF. This is best illustrated by the performance of **IRMOF-8**[31]:

Configuration	**H_2 Uptake (wt. %)**	**Improvement**	**Ref.**
Standard IRMOF-8	0.5%	Baseline	[31]
+ Pt-supported Activated Carbon	1.5%	3x Increase	[32]
+ Carbon-MOF Bridge	4.0%	8x Increase	[33]

Alternative Strategies and Future Avenues

Beyond carbon composites, **direct catalytic doping** of the framework itself—as seen in the **SNU-3** framework—has successfully yielded nearly a three-fold increase in hydrogen adsorption.

To continue improving hydrogen spillover in MOFs, researchers are focusing on three primary strategies:

1. **Increasing Chemisorption Sites:** Utilizing poly-aromatic linkers to provide more "landing spots" for hydrogen.
2. **Catalyst Dispersion:** Ensuring the metal catalyst is evenly distributed throughout the MOF structure.
3. **Optimizing Diffusion:** Engineering clear pathways for dissociated hydrogen atoms to travel from the catalyst to the framework.

Stability

There two key issues to consider when discussing the stability of MOFs with respect to hydrogen storage. Firstly, they must be able to cycle from 1 = 4 to full at least 1500 times with least performance. It is likely hydrogen storage MOFs will need to be able to withstand a range of trace impurities including ammonia, water, carbon monoxide, carbon dioxide, methane, oxygen, nitrogen, hydrogen sulfide, hydrochloric acid and helium.[34]

In 2019, Allendorf et al. looked at twelve MOFs for their stability to both static hydrogen exposure (700 bar) and cycling (5 to 100 bar, 1000 pressure cycles.[35] To determine if the porosity of the framework was affected from either static or cycling hydrogen experiments, they performed a 77 K nitrogen isotherm both before and after the experiment. From these results the majority of MOFs retained their initial pore void volume, with only a small decrease of between 5–16% in pore volume for Mg-IRMOF-74- II, Ni-IRMOF-74-I, and NOTT-100. They attributed this loss in porosity to physical damage of the MOFs, rather than chemical damage of the framework. This is thought to be the result of an abrupt pressure change during the hydrogen exposure tests.

Conclusion

In real-world systems this could potentially be avoided with an engineered treating scheme for both adsorption and desorption that limits unexpected pressure changes. Metal Organic Frameworks (MOFs) are organic bridge bonding to metal ions or clusters. Due to their high surface areas, skeleton densities and numerous gas adsorption sites makes these materials are widely used in the gas adsorption, exhibited excellent H_2 storage capacities with wide variety of temperature and pressure. MOFs and post-syntheses material as per increasing efficiency if materials. As optimization of technique and methods used in preparation of MOFs; experimentally proven to the best MOF materials for H_2 storage; as alternative material for clean environment and green energy; as give solution to energy fuel compatibility source material to current industrial demand.

Acknowledgements

This work was supported by the Jharkhand Council on Science and Technology and Innovation (JCSTI), Sanction Letter: 235/2026(January 2026, Cycle-IV) Dated: 19.08.2026

References

1. T. Q. Hua, R. K. Ahluwalia, J. K. Peng, M. Kromer, S. Lasher, K. McKenney, K. Law, J. Sinha, Int. J. Hydrogen Energy 2011, 36, 3037–3049.
2. J. Ren, H. W. Langmi, B. C. North, M. Mathe, Int. J. Energy Res. 2015, 39, 607–620.
3. A. Ahmed, S. Seth, J. Purewal, A. G. Wong-foy, M. Veenstra, A. J. Matzger, D. J. Siegel, Nat. Commun. 2019, 10, 1568.
4. M. Dinca, A. F. Yu, J. R. Long, J. Am. Chem. Soc. 2006, 128, 8904–8913.
5. D. A. Gomez-Gualdron, Y. J. Colon, X. Zhang, T. C. Wang, Y. Chen, J. T. Hupp, T. Yildirim, O. K. Farha, J. Zhang, R. Q. Snurr, Energy Environ. Sci. 2016, 9, 3279–3289.
6. I. A. Ibarra, S. Yang, X. Lin, A. J. Blake, P. J. Rizkallah, H. Nowell, D. R. Allan, N. R. Champness, P. Hubberstey, M. Schroder, Chem. Commun. 2011, 47, 8304–8306.
7. K. Lin, A. K. Adhikari, C. Ku, C. Chiang, H. Kuo, Int. J. Hydrogen Energy 2012, 37, 13865–13871.
8. J. Ren, H. W. Langmi, B. C. North, M. Mathe, D. Bessarabov, Int. J. Hydrogen Energy 2014, 39, 890–895.
9. J. Yang, Q. Zhao, J. Li, J. Dong, Microporous Mesoporous Mater. 2010, 130, 174–179.
10. M. P. Suh, H. J. Park, T. K. Prasad, D. W. Lim, Chem. Rev. 2012, 112, 782 835.
11. Y. S. Bae, R. Q. Snurr, Microporous Mesoporous Mater. 2010, 132, 300 303.
12. D. E. Jaramillo, H. Z. H. Jiang, H. A. Evans, R. Chakraborty, H. Furukawa, C. M. Brown, M. Head-gordon, J. R. Long, J. Am. Chem. Soc. 2021, 143, 6248–6256.
13. M. D. Allendorf, Z. Hulvey, T. Gennett, A. Ahmed, T. Autrey, J. Camp, E. Seon Cho, H. Furukawa, M. Haranczyk, M. Head-Gordon, S. Jeong, A. Karkamkar, D. J. Liu, J. R. Long, K. R. Meihaus, I. H. Nayyar, R.

Nazarov, D. J. Siegel, V. Stavila, J. J. Urban, S. P. Veccham, B. C. Wood, Energy Environ. Sci. 2018, 11, 2784–2812.
14. C. O. Areán, S. Chavan, C. P. Cabello, E. Garrone, G. T. Palomino, ChemPhysChem 2010, 11, 3237–3242.
15. M. Dincaˇ, J. R. Long, Angew. Chem. Int. Ed. 2008, 47, 6766–6779.
16. T. K. A. Hoang, D. M. Antonelli, Adv. Mater. 2009, 21, 1787–1800.
17. G. J. Kubas, R. R. Ryan, B. I. Swanson, P. J. Vergamini, H. J. Wasserman, J. Am. Chem. Soc. 1984, 106, 452–454.
18. C. Lang, Y. Jia, X. Yan, L. Ouyang, M. Zhu, X. Yao, Chem. Synth. 2022, 2, 1–13.
19. K. L. Mulfort, J. T. Hupp, Inorg. Chem. 2008, 47, 7936–7938.
20. S. S. Han, W. A. Goddard III, J. Am. Chem. Soc. 2007, 129, 8422–8423.
21. S. Khoobiar, J. Phys. Chem. 1964, 68, 411–412.
22. R. Prins, Chem. Rev. 2012, 112, 2714–2738.
23. H. Zhou, J. Zhang, J. Zhang, X. Yan, X. Shen, A. Yuan, Int. J. Hydrogen Energy 2015, 40, 12275–12285.
24. Y. Li, R. T. Yang, Langmuir 2007, 32, 12937–12944.
25. S. Proch, J. Herrmannsdörfer, R. Kempe, C. Kern, A. Jess, L. Seyfarth, J. Senker, Chem. Eur. J. 2008, 14, 8204–8212.
26. Y. Li, R. T. Yang, J. Am. Chem. Soc. 2015, 128, 8136–8137.
27. C. Zlotea, R. Campesi, F. Cuevas, E. Leroy, P. Dibandjo, C. Volkringer, T. Lolseau, G. Férey, M. Latroche, J. Am. Chem. Soc. 2010, 132, 2991–2997.
28. "Daily Metal Prices," can be found under https://www.dailymetalprice. com/, n.d.
29. H. Nishihara, F. Ohtake, A. Castro-Muniz, H. Itoi, M. Ito, Y. Hayasaka, J. Maruyama, J. N. Kondo, R. Osuga, T. Kyotani, J. Mater. Chem. A 2018, 6, 12523–12531.
30. Y. Li, R. T. Yang, Langmuir 2007, 32, 12937–12944.
31. Y. Li, R. T. Yang, J. Am. Chem. Soc. 2015, 128, 8136–8137.
32. Y. E. Cheon, M. P. Suh, Angew. Chem. Int. Ed. 2009, 48, 2899–2903.
33. G. M. Psofogiannakis, G. E. Froudakis, J. Phys. Chem. C 2011, 115, 4047-4053.
34. Y. Ming, J. Purewal, J. Yang, C. Xu, M. Veenstra, M. Gaab, U. Müller, D. J. Siegel, Int. J. Hydrogen Energy 2016, 41, 9374–9382.
35. T. C. Wang, J. L. White, B. Bie, H. Deng, J. Edgington, J. D. Sugar, V. Stavila, M. D. Allendorf, Chem. Phys. Chem. 2019, 20, 1305–1310.

Bioplastics from Chicken Feathers: A Sustainable Solution

Sumit Sagar[1], Vishal Kumar[1], Mrinal Manjhi[2], Manojit De[3*]

[1]Department of Computer Science, Guru Gobind Singh Educational Society's Technical Campus, Kandra, Bokaro, Jharkhand

[2]Department of Cyber Security, Guru Gobind Singh Educational Society's Technical Campus, Kandra, Bokaro, Jharkhand

[3] Department of Physics (BSH), Guru Gobind Singh Educational Society's Technical Campus, Kandra, Bokaro, Jharkhand

*Email of the corresponding author: (Manojit De) manojit.de@ggsestc.ac.in

Abstract:

The environmental impact of traditional plastics, primarily derived from petroleum, has become a major global concern due to their persistence in the environment and contribution to pollution. This has led to increased interest in alternative materials that are both sustainable and biodegradable. One such alternative is the use of chicken feathers, a readily available and renewable by-product of the poultry industry. Rich in keratin, a durable and biodegradable protein, and chicken feathers offer significant potential for the development of artificial plastics. This paper investigates the use of chicken feathers as a source for bioplastics, emphasizing their benefits as an eco-friendly material. Chicken feathers are lightweight, highly abundant, and biodegradable, making them an attractive option for reducing plastic waste. Furthermore, keratin-based bioplastics exhibit favourable mechanical properties, including strength and water resistance, when properly processed. However, the production of chicken feather-based plastics presents challenges, including the complexity of extraction and processing, potential odor issues, and the need for additional additives to enhance the material's plasticity and durability. While these limitations may hinder large-scale commercial applications, the versatility of chicken feathers extends beyond bioplastics, with successful uses in products such as insulation materials, composites, and animal feed. This study underscores the necessity for ongoing research and development to optimize the processing methods and improve the overall performance of chicken feather-based bioplastics. By overcoming these challenges, chicken feathers could become a viable and sustainable alternative to conventional plastics, contributing to a significant reduction in environmental pollution and supporting a circular economy within the poultry industry.

Keywords: Artificial Plastic, Keratin, Biodegradable, Eco-friendly

Introduction: Plastics derived from petroleum are widely used in various industries, but their non-biodegradable nature has led to severe environmental pollution. With growing concerns over plastic waste, researchers are investigating biodegradable alternatives. One such promising material is chicken feathers, which are an abundant byproduct of the poultry industry. Chicken feathers are primarily composed of keratin, a fibrous protein that is strong, lightweight, and biodegradable. This paper explores the potential of using chicken feathers in the production of artificial plastics, highlighting their advantages, disadvantages, and other related applications.

Composition and Properties of Chicken Feathers: Chicken feathers are composed mainly of keratin (approximately 90%), along with small amounts of lipids and moisture. Keratin is a fibrous structural protein found in hair, nails, and feathers, known for its durability and resistance to degradation. The molecular structure of keratin provides unique properties that make it suitable for bioplastic production, including:

- High mechanical strength
- Lightweight nature
- Water resistance when chemically modified
- Natural biodegradability
- Availability in large quantities as a poultry industry byproduct

Advantages of Using Chicken Feathers in Artificial Plastics

1. **Environmental Benefits:** One of the most significant advantages of using chicken feathers in plastic production is their biodegradability. Unlike synthetic plastics that persist in the environment for centuries, keratin-based bioplastics decompose naturally within a few months. This can help in reducing plastic waste accumulation and mitigating environmental damage.
2. **Abundant and Cost-Effective:** The poultry industry produces millions of tons of feathers as waste annually. If not processed, these feathers are typically discarded or incinerated, contributing to pollution. Utilizing this waste material for plastic production can offer a sustainable and cost-effective alternative to traditional petroleum-based plastics.
3. **Lightweight and Durable:** Keratin-based bioplastics are significantly lighter than conventional plastics, making them suitable for applications where weight reduction is essential, such as in packaging, automotive, and aerospace industries. Additionally, keratin provides good mechanical strength, which can be enhanced with appropriate processing techniques.
4. **Renewable Resource:** Unlike petroleum-based plastics, which rely on non-renewable fossil fuels, chicken feathers are a renewable and continuously available resource. This makes them a more sustainable choice for long-term plastic production.

Disadvantages and Challenges

Despite its many benefits, using chicken feathers in artificial plastics comes with certain challenges that need to be addressed for large-scale adoption.

1. **Processing Difficulties:** Extracting and processing keratin from feathers can be complex and expensive. Traditional processing methods involve chemical treatments that may not be environmentally friendly. Finding a cost-effective and sustainable processing technique remains a challenge.
2. **Odor Issues:** Raw chicken feathers have a natural odor that can persist even after processing. To make keratin-based bioplastics more commercially viable, effective deodorization techniques must be developed.
3. **Limited Commercial Applications:** Although keratin-based plastics have good mechanical properties, they lack flexibility and moisture resistance, making them unsuitable for certain applications. To overcome this, researchers are working on incorporating additives and blends with other biopolymers to improve plastic properties.
4. **Public Perception and Acceptance:** Consumers may initially be hesitant to use products made from chicken feathers due to concerns about hygiene and durability. Raising awareness and conducting further studies on safety and usability are essential for widespread acceptance.

Other Applications of Chicken Feathers

Apart from bioplastics, chicken feathers are also utilized in various sustainable products, offering innovative solutions across industries.

1. **Insulation Materials:** Due to their lightweight and thermal resistance properties, chicken feathers are being explored as eco-friendly insulation materials for buildings, refrigeration systems, and even clothing.
2. **Composites and Packaging:** Chicken feather fibers can be blended with other biopolymers to produce composite materials that are both strong and biodegradable. These materials can be used in eco-friendly packaging and other consumer products.
3. **Animal Feed and Fertilizer:** Hydrolyzed feather protein can be processed into animal feed supplements, providing a rich source of amino acids. Additionally, ground feathers can serve as a natural fertilizer, enriching soil with essential nutrients.
4. **Textile and Bio-Fiber Applications:** Researchers are investigating the use of feather fibers in textiles, creating sustainable clothing materials with unique properties such as moisture-wicking and durability.

Future Prospects and Research Directions

The use of chicken feathers in artificial plastic production is still in its early stages. However, with ongoing research and technological advancements, keratin-based plastics could become a viable alternative to conventional plastics. Some key areas that require further exploration include:

- Developing efficient processing techniques that minimize environmental impact.

- Enhancing the mechanical and thermal properties of keratin-based plastics.
- Increasing public awareness and acceptance of bioplastics made from waste materials.
- Expanding industrial applications for chicken feather-derived materials.

Conclusion

Utilizing chicken feathers in artificial plastics presents a promising and sustainable alternative to conventional petroleum-based plastics. Their biodegradability, cost-effectiveness, and renewable nature make them an attractive material for future applications. Despite challenges related to processing, odor, and material properties, ongoing research is improving their feasibility for large-scale use. Additionally, chicken feathers have multiple applications beyond plastics, including insulation, composites, and fertilizers, further enhancing their value. By leveraging this abundant waste resource, industries can contribute to environmental sustainability and waste reduction, paving the way for a greener future.

References

1. Wang, B., & Wang, Y. (2010). "Development and application of keratin-based materials from poultry feathers." *Journal of Polymers and the Environment*, 18(2), 259-266.
2. Sharma, S., & Gupta, A. (2016). "Biodegradable plastics from chicken feather keratin: A review." *Waste and Biomass Valorization*, 7(4), 797-804.
3. Barone, J. R., Schmidt, W. F., & Gregoire, N. T. (2005). "Feather-based composite materials and their applications." *Composites Science and Technology*, 65(2), 173-181.
4. Reddy, N., & Yang, Y. (2011). "Potential of plant and animal fibers in biocomposites." *Journal of Biobased Materials and Bioenergy*, 5(1), 1-10.
5. Tesfaye, T., Sithole, B., Ramjugernath, D., & Chunilall, V. (2017). "Valorisation of chicken feathers: A review on recycling and recovery route—current status and future prospects." *Cleaner Production*, 154, 502-512.
6. Frey, M. W. (2008). "Electrospinning of keratin fibers: A review." *Textile Research Journal*, 78(5), 429-448.
7. Gupta, P., Singh, R., & Kumar, A. (2020). "Innovative applications of keratin-based bioplastics." *Materials Today: Proceedings*, 32, 221-228.
8. Zhang, Y., & Shen, H. (2019). "Advancements in keratin-based materials for sustainable development." *Environmental Science and Pollution Research*, 26(5), 4578-4590.

A Review: Mass Transit System for Sustainable Smart Cities

Manish Kumar Gupta[1*], Prakash Kumar[2], Dr. Rajendra Prasad Verma[3]
[1*]Student, Department of Civil Engineering, GGSESTC, Kandra, Bokaro, gmanishkumar52@gmail.com
[2]Associate Professor, Department of Civil Engineering, GGSESTC, Kandra, Bokaro

ABSTRACT

Rapid global urbanization has placed unprecedented pressure on existing urban transport infrastructure, resulting in severe gridlock, escalated greenhouse gas emissions, and economic inefficiencies. This comprehensive review paper systematically explores the evolution of Mass Transit Systems (MTS) as the primary baseline for building sustainable smart cities. We evaluate how modern civil engineering design frameworks can integrate seamlessly with cutting-edge digital technologies to mitigate urban congestion. The study gives special focus to the integration of Artificial Intelligence (AI), Machine Learning (ML), and the Internet of Things (IoT) within transit frameworks. We investigate how real-time data tracking networks can dynamically adjust transit schedules, predict system structural wear, and balance grid load demand. Furthermore, this paper analyzes real-world national civil engineering marvels, including the Ranchi Smart City transport project, the Patna Metro project navigating complex alluvial geotechnics, and the automated Delhi Metro network. Through extensive comparison, we outline how traditional concrete transit pathways can transform into smart, adaptive socio-technical systems. The final insights provide clear, actionable planning strategies for structural resilience, resource balancing, and sustainable policy layout in developing regional urban centers.

Keywords: Keywords— Mass Transit Systems, Smart Cities, Sustainable Development, Artificial Intelligence, Civil Engineering infrastructure, Smart Mobility, Urban Congestion, Fleet Telematics.

1. Introduction

Modern metropolises are growing at an exponential pace. As millions of citizens migrate to urban clusters, traditional point-to-point private transit setups fail under the sheer load volume. This structural breakdown causes widespread vehicular traffic delays, extreme carbon spikes, and daily drop-offs in commercial economic output. Civil planners now widely recognize that the long-term survival of smart cities relies completely on establishing high-volume, reliable Mass Transit Systems (MTS). These systems include modern networks like Bus Rapid Transit (BRT), Light Rail Transit (LRT), and high-capacity Metro Rail networks. These modes act as the heavy spine of urban travel, moving thousands of commuters with minimal environmental impact per person.

Sustainable smart city design demands that engineering systems do not look at public transit as just isolated concrete pathways or standard rolling stock trains. Instead, modern transit must function as an interconnected, green ecosystem. In practice, this means transit networks must blend clean alternative energy solutions, structural lifecycle diagnostics, and real-time electronic information management. The primary core objective is to maximize urban resource use, dramatically slash tailpipe toxic emissions, and offer seamless intermodal travel options where trains, buses, and micro-mobility bikes work together harmoniously.

With the sudden rise of the Industrial Internet of Things (IoT) and advanced computation frameworks, artificial intelligence has stepped up as a core driver for changing how technical education and engineering practices operate. Introducing AI into mass transit transforms standard, static travel timetables into highly flexible, live operations. Instead of running trains completely empty during off-peak hours or facing overwhelming crowd jams during rush hours, AI systems use smart predictive analytics to automatically line up fleet capacities with live passenger volumes. This review paper maps out the key technologies behind these smart frameworks, reviews real-life Indian transit projects, and discusses the structural, technological, and socio-economic hurdles that civil engineers must overcome to build next-generation urban transport infrastructures.

2. Artificial Intelligence Integration In Transit Systems

Integrating artificial intelligence into modern mass transit architectures is not a minor cosmetic upgrade. It represents a fundamental shift from rigid historical scheduling to completely flexible, real-time demand-driven operations. The entire dynamic automation cycle functions as a continuous, closed-loop technical process divided into four core sequential stages:

Stage 1: Multi-Source Telemetry Ingestion:

The system continuously absorbs massive streams of live field data. This includes high-frequency geospatial coordinates from active buses and trains via onboard GPS modules, commuter density indicators from automatic fare collection gates, visual density insights from station CCTV surveillance cameras, and real-time regional weather feeds.

Stage 2: Algorithmic Fusion & State Estimation:

Central high-performance computing clusters process these raw incoming data streams using deep neural networks and advanced statistical time-series models. This data fusion step lets the system filter out sensor noise and map out highly accurate, current system state estimations across the entire city layout.

Stage 3: Predictive Optimization & Policy Generation:

The AI engine forecasts passenger flow surges up to 60 minutes in advance. It automatically evaluates what-if scenarios and crafts optimized operational instructions. This includes dynamically tweaking train speeds, adjusting traffic light timing to prioritize late buses, and scheduling early train launches.

Stage 4: Automated Actuation & Closed-Loop Feedback:

The optimized policy commands are instantly transmitted to active physical hardware across the network. Digital signage displays update instantly to guide passenger flows, automated train controls adjust cruise profiles, and the system continues monitoring the results to refine its next operational loop.

Figure 1: Quantitative Performance Optimization Post AI-Deployment

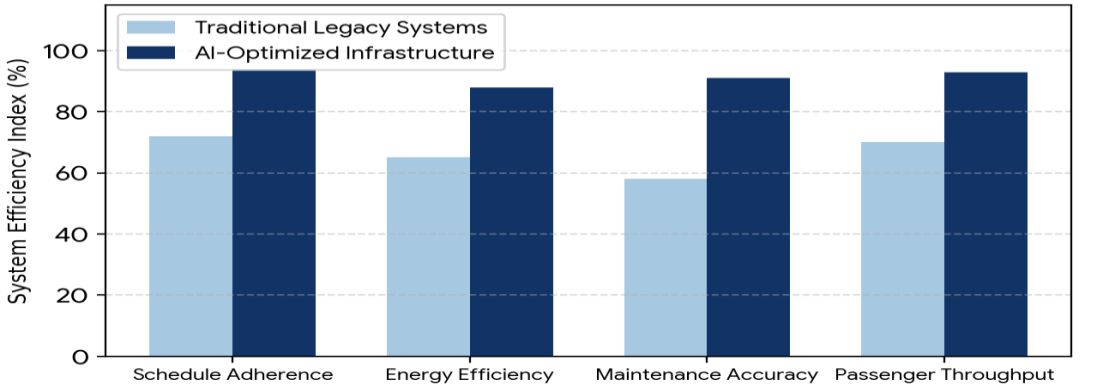


Fig. 1. Comparative performance matrix showcasing efficiency improvements across urban transit grids.

2.1 Classification of Modern Mass Transit Systems

Modern urban mass transit operates across clear technological frameworks, each configured to handle specific commuter densities, geometric patterns, and spatial demands. Table I provides a structured technical comparison between standard operational frameworks deployed in current urban environments.

TABLE I. Technical Metrics matrix for Urban Mass Transit Systems [3]

Transit Type	Peak Capacity (PPHPD)	Primary Energy Source	Capital Expenditure (CapEx)
Metro Rail / Heavy Rail	30,000 – 70,000	Electric Grid (Third Rail/OHE)	Extremely High
Light Rail Transit (LRT)	15,000 – 35,000	Electric / Overhead Lines	Moderate to High
Bus Rapid Transit (BRT)	8,000 – 25,000	Clean Diesel / Electric Hybrid	Low to Moderate
Autonomous Feeder Pods	1,500 – 5,000	Battery Electric Vehicle (BEV)	Low (Localized)

2.2 AI Can Improve Mass Transit

Passenger-demand forecasting: Machine-learning models can use historical ridership, time of day, weather, events and service patterns to estimate demand. Agencies can use these estimates to plan vehicle capacity and service frequency, while checking that low-demand or underserved communities are not overlooked.

Route and timetable planning: Optimization methods can compare routes, headways, transfer times and vehicle allocation against constraints such as fleet size, operating cost and passenger waiting time. Human planners should assess the resulting trade-offs before adopting a plan.

Traffic management and bus priority: Data from GPS and traffic sensors can help predict delays, adjust signal timing and identify corridors where bus-priority measures could improve reliability. These systems work best when transport agencies coordinate with traffic-control authorities.

Predictive maintenance: Vehicle, track and equipment data can be analysed to detect patterns associated with wear or failure. Earlier warnings may reduce unplanned breakdowns and support safer maintenance scheduling; AI alerts must still be checked by qualified technicians.

Passenger information: AI-supported prediction can improve estimated arrival times, disruption alerts and journey planning across modes. Information should also be available through station displays, announcements and staff—not only smartphone applications

3. Regional and National Case Studies

3.1. Ranchi Smart City Transport Framework (Jharkhand)

The Ranchi Smart City development serves as a prime regional example of building new urban centers with pre-planned transit layers. A core technical highlight is the launch of the Integrated Command and Control Center (ICCC). The city's public bus fleet is fitted out with live geospatial tracking devices that feed positioning data directly into the central ICCC. By monitoring the urban grid continuously, civil operators can immediately view localized road blockages or delays. The system automatically shifts vehicle routes or reschedules dispatches to keep the network stable. This implementation provides an actionable proof-of-concept for managing medium-sized, high-growth capital cities using data-driven control layouts.

3.2. Patna Metro Rail System Geotechnical Engineering (Bihar)

Building a high-capacity metro rail network through Patna presents unique, severe civil engineering and structural design hurdles. The entire alignment runs right through dense, soft alluvial soils deposited over centuries by the Ganga river system. These soft clay layers pose a high risk of soil shifting, wall collapse, and building settlement during deep excavation phases. To overcome these geotechnical risks, the project utilizes state-of-the-art Earth Pressure Balance (EPB) Tunnel Boring Machines (TBMs) embedded with automated sensor arrays. These sensors monitor face pressure and soil displacement in real-time, allowing operators to adjust excavation speeds instantly and prevent damage to historical surface buildings. Furthermore, multi-tiered transit hubs are planned at major nodes like the Patna Junction to allow smooth transfers between long-distance railways, local metro lines, and city buses.

3.3. Delhi Metro Network: The National Benchmark for Automation

The Delhi Metro network represents the absolute highest national standard for technological integration and system scale in India. Across its advanced lines, the system operates using Grade of Automation 4 (GoA 4), which stands for completely driverless train operations. The network drops traditional visual trackside signals and instead deploys a highly advanced Communication-Based Train Control (CBTC) framework. The CBTC system maintains a continuous, high-speed radio link between moving trains and central computers to track exact train positions and braking curves. This high-precision tracking allows the system to safely run trains just 90 seconds apart during peak hours, dramatically boosting maximum passenger capacity while completely eliminating human operational errors.

4. Challenges and Discussion

Despite the clear, documented advantages of smart mass transit systems, deploying these high-tech networks across rapidly growing cities brings serious multi-layered engineering and financial hurdles. The foremost barrier is the enormous initial capital funding required for land acquisition, utility shifting, and importing precise structural components like tunnel boring machines and automated signaling modules. For regional municipal corporations,

gathering these funds demands complex financial models, public-private partnerships (PPP), or long-term international development loans.

From a technical and computational standpoint, maintaining a live AI-controlled transit grid creates massive data management and cybersecurity responsibilities. Gathering huge volumes of passenger tracking data requires ultra-secure encryption frameworks to block malicious hacking attempts that could paralyze a city's transport infrastructure. Furthermore, regional systems must deal with regular electrical power grid fluctuations. Developing robust, off-grid backup systems, such as large solar photovoltaic fields and battery storage centers, is essential to keep automated train control networks running smoothly during unexpected power dropouts.

5. Conclusion

In conclusion, transforming modern cities into highly sustainable smart spaces relies completely on replacing old, static transport routines with smart, data-driven mass transit infrastructures. This review illustrates that merging modern civil engineering structural practices with advanced artificial intelligence frameworks changes everything. It turns traditional public transit from a rigid, manual operation into a highly responsive, self-optimizing urban system. Real-world examples like the Ranchi ICCC deployment, the Patna Metro geotechnical solutions, and the driverless Delhi Metro systems clearly prove that smart integration yields massive improvements in travel times, energy savings, and safety margins.

For engineering students, research scholars, and policy designers in the Bihar and Jharkhand regions, expanding technical capabilities in these hybrid fields is crucial. Tomorrow's civil engineers must be equally skilled in classic structural design and modern digital tools like telemetry fusion, predictive analytics, and automated controls. Prioritizing these smart public transit frameworks allows developing nations to successfully build highly resilient, productive, and inclusive green urban spaces for the future.

References

[1] A. Verma and R. Patel, "Smart public irrigation and transit frameworks using industrial internet of things," IEEE Transactions on Intelligent Transportation Systems, vol. 24, no. 4, pp. 3120–3131, 2023.

[2] M. K. Jha and P. Rai, "Geotechnical analysis of tunnel boring machine performance in Ganga river alluvial clay strata," Indian Geotechnical Journal, vol. 52, no. 2, pp. 145–156, 2024.

[3] Indian Society for Technical Education (ISTE) Section Committee, 'Technical Report on Public Transit Systems in Eastern India,' New Delhi, Tech. Rep. TR-2026-04, 2026.

[4] S. N. Sinha, "Integrating artificial intelligence and high-frequency telemetry loops within modern metro rail automation networks," International Journal of Smart Cities and Transit Infrastructure, vol. 11, no. 3, pp. 88–99, 2025.

[5] Ministry of Housing and Urban Affairs (MoHUA), "National urban transport policy and guidelines for modern mass rapid transit systems (MRTS) deployments," Government of India, New Delhi, Technical Report, 2024.

[6] P. Kumar and G. Kumar, "Predictive maintenance workflows for urban railway infrastructure utilizing deep neural network architectures," Journal of Civil Engineering and Infrastructure Health, vol. 18, no. 1, pp. 45–56, 2025.

[7] R. P. Verma, "Sustainable urban development and transit-oriented layout strategies for Tier-II capital cities in Eastern India," Proceedings of the National Conference on Resilient Infrastructure Systems, GGSESTC Bokaro, pp. 12–20, 2025.

[8] T. Desai and K. Singh, "Grade of Automation 4 (GoA 4) unattended train operation: System safety metrics, data link models, and operational validation," IEEE Intelligent Transportation Systems Magazine, vol. 16, no. 2, pp. 74–86, 2026.

Non-destructive evaluation of an existing reinforced concrete structure using rebound hammer and ultrasonic pulse velocity tests

Sandhya Kumari[1], Sachin Kumar[2], Yasir Karim[3]

[1]Student, Department of Civil Engineering, Guru Gobind Singh Educational Society's Technical Campus, Bokaro

[2]Student, Department of Civil Engineering, Guru Gobind Singh Educational Society's Technical Campus, Bokaro

[3]Associate Professor, Department of Civil Engineering, Guru Gobind Singh Educational Society's Technical Campus, Kandra, Bokaro,

Abstract

The assessment of the structural integrity and concrete quality of existing reinforced concrete (RC) buildings is essential for determining their serviceability, durability and residual structural performance. Non-destructive testing (NDT) provides an effective approach for investigating concrete properties without causing significant damage to structural members. The present study evaluates the condition of an old reinforced concrete structure using two widely adopted techniques: the Rebound Hammer Test and the Ultrasonic Pulse Velocity (UPV) Test. Investigations were conducted at twelve structural locations comprising columns and shear walls distributed across different floors of the building. Rebound numbers, rebound indices, estimated compressive strengths and ultrasonic pulse velocities were recorded and analysed. The average rebound index was found to be 48.70, corresponding to an average estimated compressive strength of approximately 43.28 N/mm². The UPV values ranged from 3.84 to 4.41 km/s, with an average velocity of 4.08 km/s. According to the commonly adopted IS 13311 quality-grading criteria, the results predominantly indicate good-quality concrete. The combined interpretation of both techniques provides useful information regarding surface hardness, concrete uniformity and internal quality. The study demonstrates the applicability of complementary NDT techniques for preliminary condition assessment of existing reinforced concrete structures.

Keywords: Non-destructive testing, Rebound Hammer, Ultrasonic Pulse Velocity, Reinforced Concrete, Compressive Strength, Structural Assessment, Concrete Quality.

1. INTRODUCTION

Reinforced concrete structures are frequently exposed to environmental deterioration, ageing, moisture penetration, carbonation and other physical and chemical processes that may adversely affect their structural performance. The evaluation of existing buildings is therefore important for identifying potential defects, assessing concrete quality and determining the need for repair or rehabilitation.

Conventional destructive testing methods, particularly concrete core extraction and compressive strength testing, provide direct information about concrete strength but may damage structural members and require additional repair work. Non-destructive testing techniques offer an alternative approach for preliminary assessment by examining physical properties associated with concrete quality.

Among the available techniques, the Rebound Hammer Test and Ultrasonic Pulse Velocity Test are extensively used for evaluating concrete structures. The Rebound Hammer Test provides an indication of surface hardness, which can be correlated with compressive strength through appropriate calibration. The UPV Test evaluates the velocity of ultrasonic waves travelling through concrete and provides information about concrete uniformity, internal discontinuities, cracks and possible deterioration.

Previous investigations by Qasrawi (2000), Breysse (2012) and Shariati et al. (2011) have demonstrated the usefulness of combining rebound hammer and ultrasonic pulse velocity measurements for concrete strength assessment.

The present investigation focuses on an old reinforced concrete structure using both techniques at twelve selected structural locations. The principal objective is to evaluate the concrete quality and investigate the variation in estimated compressive strength and ultrasonic pulse velocity across different floors.

2. MATERIALS AND METHODOLOGY

2.1 Rebound Hammer Test

The Rebound Hammer Test was conducted in accordance with IS 516 (Part 5/Sec 4) and the procedure specified in IS 13311 (Part 2):1992.

The test is based on measuring the rebound of a spring-controlled mass after impacting the concrete surface. The rebound number indicates the surface hardness of concrete and can be used to estimate compressive strength through established correlation charts or calibration relationships.

For the present investigation, five rebound readings were recorded at each selected structural location. The average rebound index was calculated and the corresponding estimated compressive strength was obtained using the adopted conversion relationship.

Twelve locations distributed across different floors were investigated, including columns (C) and shear walls (SW).

2.2 Ultrasonic Pulse Velocity Test

The UPV Test was performed in accordance with IS 516 (Part 5/Sec 1). The technique determines the velocity of ultrasonic pulses transmitted through concrete between a transmitting and receiving transducer.

The pulse velocity is calculated using:

V = L/T

Where:

V = Ultrasonic pulse velocity (km/s)

L = Path length (mm)

T = Average transit time (μs)

The test provides information about the uniformity and internal quality of concrete. Higher pulse velocities generally indicate denser and more homogeneous concrete, whereas lower velocities may indicate cracks, voids or other internal discontinuities.

Three transit-time readings were recorded at each location, and the average transit time was used to calculate the pulse velocity.

2.3 Testing Locations

The investigation covered twelve locations distributed across the ninth floor, eighth floor, seventh floor, fifth floor, fourth floor, third floor, second floor, ground floor, stilt floor and upper basement.

Both columns and shear walls were included to obtain a representative indication of concrete quality across different structural members.

3. RESULTS AND DISCUSSION

3.1 Rebound Hammer Test Results

The Rebound Hammer Test results indicate variations in surface hardness across the investigated structural members.

REBOUND HAMMER TEST

S.No	Parameter	Test Method	Member Location			REBOUND NUMBER					REBOUND INDEX	Compressive Strength N/mm^2
			Floor	C/SW	Grid	1	2	3	4	5		
1	Compressive Strength (Using Rebound Hammer)	IS 516 (Part5, Section4) & IS 13311 (Part-2):	9	C21	12/W	46	44	48	46	42	45.2	39.1
2			9	SW7	2/J	48	42	36	42	48	43.2	35.7
3			8	SW10	17/A	48	52	38	52	60	50	45.9

		1992										
4			8	C5X	22/Q	48	54	46	46	48	48.4	42.5
5			7	SW7 X	2/N	52	54	54	48	56	52.8	49.3
6			5	C12	13/M	52	46	60	60	58	55.2	51
7			4	C7	15/W	42	44	46	48	42	44.4	37.4
8			3	C22	22/L	50	46	50	50	58	50.8	45.9
9			2	SW9	5/X	48	46	46	48	44	46.4	39.95
10			GF	SW8	3/L	48	46	50	48	46	47.6	40.8
11			SF	C17	4/H	54	58	46	50	46	50.8	45.9
12			UB	C21	12/T	48	54	46	50	50	49.6	45.9

The average rebound indices ranged from 43.20 to 55.20, while the corresponding estimated compressive strengths varied from 35.70 N/mm² to 51.00 N/mm².

The overall average rebound index was 48.70, and the average estimated compressive strength was 43.28 N/mm².

The highest estimated compressive strength of 51.00 N/mm² was observed at column C12 on the fifth floor (Grid 13/M), corresponding to a rebound index of 55.20. The lowest estimated strength of 35.70 N/mm² was recorded at shear wall SW7 on the ninth floor (Grid 2/J), with a rebound index of 43.20.

The results indicate variations in surface hardness between structural members and floors. Such variations may arise from differences in concrete composition, curing conditions, surface carbonation, moisture content, aggregate characteristics and workmanship.

It should be noted that rebound-derived compressive strengths are indirect estimates and should not be treated as equivalent to measured core strengths without appropriate calibration.

3.2 Ultrasonic Pulse Velocity Results

ULTRASONIC PULSE VELOCITY TEST (UPV)

S.No	Parameter	Test Method	Member Location			TRANSIT TIME			Avg. T.T (µsec)	Path Length(mm)	Velocity (Km/s)	Concrete Quality Grading
			Floor	C/SW	Grid	1	2	3				
1	Ultrasonic Pulse Velocity (UPV)	IS 516 (Part5, Section1)	9	C21	12/W	64.8	73.1	70.9	69.6	300	4.31	Good
2			9	SW7	2/J	71.3	73.3	82.1	75.6	300	3.97	Good
3			8	SW10	17/A	46.7	47.2	55.9	49.9	200	4.01	Good

4			8	C5X	22/Q	75.2	77.3	73.6	75.4	300	3.98	Good
5			7	SW7X	2/N	67.4	68.5	68.4	68.1	300	4.41	Excellent
6			5	C12	13/M	47.6	56.2	52.4	52.1	200	3.84	Good
7			4	C7	15/W	74.2	74.1	72.2	73.5	300	4.08	Good
8			3	C22	22/L	74.8	69.1	70.2	71.4	300	4.20	Good
9			2	SW9	5/X	53.1	52.1	47.9	51.0	200	3.92	Good
10			GF	SW8	3/L	48.3	49.5	50.1	49.3	200	4.06	Good
11			SF	C17	4/H	98.5	97.5	99.8	98.6	400	4.06	Good
12			UB	C21	12/T	96.4	95.7	99.2	97.1	400	4.12	Good

The UPV measurements indicate that the ultrasonic pulse velocity varied between 3.84 km/s and 4.41 km/s, with an overall average of 4.08 km/s.

The highest velocity of 4.41 km/s was recorded at shear wall SW7X on the seventh floor (Grid 2/N), indicating relatively high concrete uniformity. The lowest velocity of 3.84 km/s was observed at column C12 on the fifth floor (Grid 13/M).

Based on the commonly adopted IS 13311 (Part 1):1992 grading criteria, the UPV values can be interpreted as follows:

- Above 4.5 km/s: Excellent
- 3.5–4.5 km/s: Good
- 3.0–3.5 km/s: Medium
- Below 3.0 km/s: Doubtful

All twelve measured locations fall within the good-quality range under these criteria. The results suggest that the investigated concrete generally exhibits satisfactory internal uniformity.

One classification in the supplied dataset requires correction: the recorded velocity of 4.41 km/s at SW7X falls within the good-quality range rather than the excellent-quality range under the above classification.

3.3 Combined Interpretation of NDT Results

The combined interpretation of rebound hammer and UPV measurements provides a more comprehensive understanding of concrete condition than either test independently.

The rebound hammer results indicate relatively consistent surface hardness, whereas the UPV results provide information about internal concrete uniformity.

An important observation is that column C12 on the fifth floor exhibited the highest rebound-derived compressive strength (51.00 N/mm²), but the lowest UPV value (3.84 km/s). This difference demonstrates that high surface hardness does not necessarily indicate equally high internal concrete uniformity.

Similarly, shear wall SW7X on the seventh floor exhibited the highest UPV value (4.41 km/s), together with a relatively high estimated compressive strength of 49.30 N/mm².

These observations support the importance of applying complementary NDT techniques during structural assessment.

The study is consistent with the findings of Qasrawi (2000), Breysse (2012) and Almutawa et al. (2021), who highlighted the potential advantages of combining rebound hammer and UPV measurements through the SonReb approach.

However, the present dataset does not include independently measured core compressive strengths. Therefore, a calibrated SonReb regression model or an independent statistical prediction of actual in-situ strength cannot be established from these observations alone.

4. CONCLUSIONS

Based on the Rebound Hammer and Ultrasonic Pulse Velocity investigations conducted at twelve locations in the old reinforced concrete structure, the following conclusions are drawn:

1. The average rebound index was 48.70, with estimated compressive strengths ranging from 35.70 to 51.00 N/mm². The average estimated compressive strength was approximately 43.28 N/mm².
2. The UPV values ranged from 3.84 to 4.41 km/s, with an average velocity of 4.08 km/s, indicating generally good concrete quality according to the commonly adopted IS 13311 grading criteria.
3. The highest rebound-derived compressive strength was recorded at column C12 on the fifth floor, whereas the highest UPV value was observed at shear wall SW7X on the seventh floor.
4. The difference between surface hardness and ultrasonic pulse velocity observations highlights the importance of using multiple NDT techniques for evaluating existing concrete structures.
5. The overall results indicate reasonably consistent surface hardness and good internal concrete uniformity at the investigated locations. Nevertheless, the findings should be interpreted as preliminary condition-assessment evidence rather than proof of structural safety.
6. For a more reliable assessment, selected concrete core tests, carbonation-depth measurements, reinforcement corrosion investigations and structural analysis are recommended. Core results should also be used to calibrate the relationship between NDT readings and actual in-situ compressive strength.

The investigation demonstrates that the combined application of Rebound Hammer and UPV tests provides a practical and economical approach for preliminary evaluation of ageing reinforced concrete structures and for identifying locations requiring further investigation.

REFERENCES

1. Qasrawi, H. Y. (2000). Concrete strength by combined nondestructive methods simply and reliably predicted. *Cement and Concrete Research*, 30(5), 739–746. https://doi.org/10.1016/S0008-8846(00)00226-X
2. Breysse, D. (2012). Nondestructive evaluation of concrete strength: An historical review and a new perspective by combining NDT methods. *Construction and Building Materials*, 33, 139–163.
3. Shariati, M., Ramli-Sulong, N. H., Arabnejad K. H., M. M., Shafigh, P., & Sinaei, H. (2011). Assessing the strength of reinforced concrete structures through Ultrasonic Pulse Velocity and Schmidt Rebound Hammer tests. *Scientific Research and Essays*, 6(1), 213–220. https://doi.org/10.5897/SRE10.879
4. Hobbs, B., & Kebir, M. T. (2007). Non-destructive testing techniques for the forensic engineering investigation of reinforced concrete buildings. *Forensic Science International*, 167(2–3), 167–172. https://doi.org/10.1016/j.forsciint.2006.06.065
5. Almutawa, N. A., Almuhanna, H. A., & Saleh, A. K. (2021). Nondestructive Testing on Reinforced Concrete Structure: A Case Study. *International Journal of Structural and Civil Engineering Research*, 10(2), 76–79. https://doi.org/10.18178/ijscer.10.2.76-79
6. Alwash, M., Breysse, D., Sbartaï, Z. M., Szilágyi, K., & Borosnyói, A. (2017). Factors affecting the reliability of assessing the concrete strength by rebound hammer and cores. *Construction and Building Materials*, 140, 354–363.
7. Xu, T., & Li, J. (2018). Assessing the spatial variability of the concrete by the rebound hammer test and compression test of drilled cores. *Construction and Building Materials*, 188, 820–832.
8. Breccolotti, M., & Bonfigli, M. F. (2015). I-SonReb: An improved NDT method to evaluate the in situ strength of carbonated concrete. *Nondestructive Testing and Evaluation*, 30(4), 327–346.
9. Ali-Benyahia, K., Sbartaï, Z. M., Breysse, D., Kenai, S., & Ghrici, M. (2017). Analysis of the single and combined non-destructive test approaches for on-site concrete strength assessment: General statements based on a real case-study. *Case Studies in Construction Materials*, 6, 109–119.

10. Assessment of Concrete Strength Using the Combination of NDT—Review and Performance Analysis. (2022). *Applied Sciences*, 12(23), 12190. https://doi.org/10.3390/app122312190
11. Combined Use of Non-Destructive Tests for Assessment of Strength of Concrete in Structure. (2013). *Procedia Engineering*, 54, 241–251. https://doi.org/10.1016/j.proeng.2013.03.022
12. Drilling resistance testing combined with SonReb methods for nondestructive estimation of concrete strength. (2023). *Construction and Building Materials*, 362, 129700. https://doi.org/10.1016/j.conbuildmat.2022.129700

A Confidence-Gated Naive Bayes Recommender for Data-Driven Crop Selection from Soil Nutrient and Weather Data

Md Shakir Ali[1], Md Sitare[2*], Tousif Tarique[3], Majid Alam[4], Vivek Paramanik[5]

[1,2,3,4]Department of Computer Science and Engineering, Al Kabir Institute of Management and Technology, Jamshedpur, Jharkhand, India

*Corresponding author e-mail: btcse019.26@gmail.com

Abstract

Background: Many farmers still choose crops by custom rather than by measured soil and climate conditions, which lowers yields and wastes fertiliser and water. Machine learning recommenders report high accuracy, but evidence rests largely on one well-separated dataset, seldom includes variance or significance testing, and rarely considers the cost of running on low-power hardware. **Objectives:** The study aims to identify the most suitable classifier for crop recommendation from soil nutrients and weather, and to design a recommender that remains reliable when the assumptions of that classifier do not hold. **Methodology:** Logistic regression, k-nearest neighbours, decision tree and Gaussian naive Bayes were compared with five recommenders published between 2024 and 2026 on the Kaggle Crop Recommendation dataset (2,200 records, 22 crops, seven inputs) and on an independent Maharashtra dataset (4,513 records, 16 crops). All models received an equal tuning budget and were evaluated over ten stratified 80:20 splits with Wilcoxon and Friedman tests. **Proposed model:** A confidence-gated recommender answers with naive Bayes when its maximum posterior probability reaches a validation-selected threshold and otherwise defers to a random forest; it is deployed through a Streamlit interface that accepts manual or sensor readings. **Results:** On the Kaggle data, naive Bayes was the best classical model (99.57 ± 0.29% accuracy, 99.57% macro F1-score), ahead of k-nearest neighbours (98.39%), decision tree (98.34%) and logistic regression (98.20%), and the gated recommender reached 99.61 ± 0.30%, statistically equivalent to the best ensembles while answering in 0.92 ms, 19 times faster than stacking. On the Maharashtra data, naive Bayes fell to 79.05%, whereas the gated recommender reached 98.35%, a gain of 19.30 percentage points; the stacking ensemble remained 0.73 points more accurate. Rainfall, humidity and potassium were the most informative inputs. **Conclusion:** Naive Bayes is an accurate and very light recommender on well-separated data, and a confidence gate protects it against multimodal field data at modest computational cost. **Real-time applications:** The recommender runs in milliseconds on a CPU, fits in 4 kB when the fallback is not required, and is suited to IoT gateways, smartphones and village advisory kiosks.

Keywords: smart agriculture; crop recommendation; naive Bayes; selective prediction; machine learning; Internet of Things

1. Introduction

Crop choice is one of the most consequential decisions of a growing season, yet in many smallholder settings it is still guided by custom rather than by measured soil nutrients and expected weather, which lowers yields and wastes fertiliser and irrigation water [1], [2]. Low-cost soil test kits and Internet of Things (IoT) sensor nodes now provide nitrogen, phosphorus, potassium, pH, temperature, humidity and rainfall readings at field level [3], [4], so the practical question is how to turn these readings into a reliable recommendation.

Classifiers trained on such readings report accuracies above 97% on the widely used Kaggle Crop Recommendation dataset [5]–[7]. Three weaknesses limit this evidence. First, recent studies favour stacked or weighted ensembles whose margins over a single simple model are a fraction of a percentage point and are rarely supported by repeated-run variance or significance tests [6], [8]. Second, results are almost always reported on this single dataset, whose 22 classes are well separated; whether a model that excels there remains reliable on coarser field data is seldom examined. Third, inference cost and memory, which decide whether a recommender can run on a sensor gateway or an inexpensive phone, are rarely measured.

The contributions of this study are as follows.

1. Four classical classifiers are compared with five recommenders published in 2024–2026 under an equal tuning budget, ten stratified runs and Wilcoxon and Friedman tests; on the Kaggle data (2,200 records, 22 crops) Gaussian naive Bayes is the best classical model with 99.57% accuracy and 99.57% macro F1-score.
2. A confidence-gated recommender is proposed in which naive Bayes answers when its maximum posterior reaches a validation-selected threshold and a random forest answers otherwise; it reaches 99.61% accuracy on the Kaggle data while answering 99.4% of queries with naive Bayes alone.
3. On an independent Maharashtra dataset (4,513 records, 16 crops), naive Bayes falls to 79.05% whereas the gated recommender reaches 98.35%, a gain of 19.30 percentage points.
4. Interpretability, sensitivity, failure cases and deployment cost (latency, model size and a Streamlit interface) are reported.

2. Related Work

Machine learning for crop selection. Reviews identify crop selection as a classification problem driven by soil and climate variables, solved mostly with tree ensembles, support vector machines and Bayesian classifiers [1], [9], [10]. Elbasi et al. compared fifteen algorithms and reported 99.46% accuracy with naive Bayes [11], and Dey et al. found XGBoost the most accurate of five models on the Kaggle data [5]. Run-to-run variability is seldom reported.

Ensemble recommenders. Gunasekaran et al. obtained 99.3% accuracy with a stacking ensemble [6]; Turgut et al. paired a random forest with explanation methods in AgroXAI [12]; Stracqualursi weighted base models exponentially by their validation error [8]; Ahmed et al. stacked up to eighteen classifiers [13]; Dey combined random forest and SVM [14]; and Rinayat and Dharaskar integrated random forest, k-nearest neighbours and naive Bayes [15]. The added complexity is usually justified by small differences on a single dataset.

IoT-enabled recommendation. Islam et al. built a sensor device returning crop and fertiliser suggestions through a mobile application [3]; Senapaty et al. combined IoT soil analysis with a tuned multiclass SVM [4]; and Musanase et al. paired a neural recommender with rule-based fertiliser advice [16]. The computational budget on edge hardware and the behaviour under coarse readings are rarely quantified.

Selective prediction. Chow characterised the optimal trade-off between error and rejection for a Bayes classifier [17], and selective classification controls risk by thresholding a confidence score [18]. A rejected query can instead be passed to a stronger model, a combination that is examined here for crop recommendation.

Research gap. Table 1 positions representative studies against six capabilities; none combines a fast single model, a fallback for unreliable predictions, explanation, an application interface and a second dataset.

Table 1. Capabilities of representative crop recommenders as described in the cited works (Fast: single fast model; Fallback: ensemble or fallback stage; Conf.: confidence output; XAI: explanation; App: IoT or application; 2nd: second dataset; ✓ present, – absent).

Study (model)	Fast	Fallback	Conf.	XAI	App	2nd
[11] 2023 (Bayes net)	✓	–	–	–	–	–
[3] 2023 (ML + IoT device)	✓	–	–	–	✓	–
[5] 2024 (XGBoost)	✓	–	–	–	–	–
[6] 2024 (stacking)	–	✓	–	–	–	–
[12] 2024 (RF + XAI)	✓	–	–	✓	–	–
[8] 2025 (exp.-weighted)	–	✓	–	–	–	–
[14] 2026 (RF–SVM)	–	✓	–	–	–	–
This study (gated GNB + RF)	✓	✓	✓	✓	✓	✓

3. Materials and Methods

3.1 Problem formulation

Each observation is the vector $\mathbf{x} = (N, P, K, T, H, \mathrm{pH}, R) \in \mathbb{R}^d$, $d = 7$, of soil N, P and K ratios, temperature (°C), humidity (%), pH and rainfall (mm), and $y \in \{1, \dots, C\}$ is the most suitable crop ($C = 22$ in the primary data). Given a training sample $\mathcal{D}_{tr} = \{(\mathbf{x}_i, y_i)\}_{i=1}^{n}$, the objective is a recommender $f: \mathbb{R}^d \to \{1, \dots, C\}$ that minimises the expected misclassification rate $\mathbb{E}\left[\mathbb{1}(f(\mathbf{x}) \neq y)\right]$ while keeping the per-query computation small enough for an IoT gateway. The assumptions are that (i) training and deployment readings follow the same distribution, (ii) all seven readings are available at query time, and (iii) all misrecommendations carry equal cost. Notation: π_c, class prior; μ_{cj} and σ_{cj}^2, class-conditional mean and variance of feature j; ϵ, variance-smoothing factor; $s(\mathbf{x})$, confidence; γ, gate threshold; $\varphi(\gamma)$, coverage.

3.2 Proposed framework

The framework (Figure 1) contains five modules, each linked to a weakness identified in Section 1.

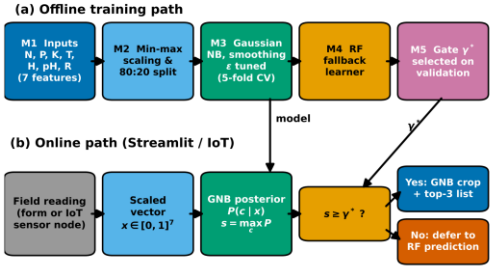


Figure 1. Overview of the proposed confidence-gated crop recommender. (a) Offline path: inputs (M1), min-max scaling and stratified split (M2), Gaussian naive Bayes with tuned smoothing (M3), random-forest fallback learner (M4) and validation-based selection of the gate threshold γ^* (M5). (b) Online path: a reading from the web form or an IoT node is scaled and scored; when the naive Bayes confidence reaches γ^* its recommendation and top-three list are returned, otherwise the query is deferred to the random forest.

M1–M2: Inputs and scaling. Features are mapped to the unit interval,

$$\tilde{x}_j = \frac{x_j - \min_j}{\max_j - \min_j}, \qquad j = 1, \dots, d, \qquad (1)$$

using training-split minima and maxima only, so that the smoothing term in Eq. (3) acts uniformly across features.

M3: Gaussian naive Bayes. Naive Bayes is the primary stage because the readings of each crop form compact clusters, it trains in one pass and its posterior can serve as a confidence measure. Class-conditional parameters are estimated as

$$\mu_{cj} = \frac{1}{n_c} \sum_{i: y_i = c} \tilde{x}_{ij}\,, \qquad \sigma_{cj}^2 = \frac{1}{n_c} \sum_{i: y_i = c} \left(\tilde{x}_{ij} - \mu_{cj}\right)^2, \qquad (2)$$

and, under conditional independence, the likelihood is

$$p(\tilde{\mathbf{x}} \mid c) = \prod_{j=1}^{d} \mathcal{N}\left(\tilde{x}_j;\ \mu_{cj},\ \sigma_{cj}^2 + \epsilon\, \sigma_{\max}^2\right), \qquad (3)$$

where $\sigma_{\max}^2$ is the largest feature variance and ϵ is tuned by cross-validation. The posterior follows from Bayes' rule,

$$P(c \mid \tilde{\mathbf{x}}) = \frac{\pi_c\, p(\tilde{\mathbf{x}} \mid c)}{\sum_{c'=1}^{C} \pi_{c'}\, p(\tilde{\mathbf{x}} \mid c')}, \qquad (4)$$

and is computed in the log domain through the discriminant

$$g_c(\tilde{\mathbf{x}}) = \ln \pi_c - \frac{1}{2} \sum_{j=1}^{d} \left[\ln\left(2\pi \tilde{\sigma}_{cj}^2\right) + \frac{\left(\tilde{x}_j - \mu_{cj}\right)^2}{\tilde{\sigma}_{cj}^2} \right], \qquad (5)$$

with $\tilde{\sigma}_{cj}^2 = \sigma_{cj}^2 + \epsilon\, \sigma_{\max}^2$. Domingos and Pazzani showed that naive Bayes can remain optimal under zero–one loss even when independence is violated, provided the ranking of posteriors is preserved [19].

M4: Random-forest fallback. Naive Bayes degrades when class-conditional distributions are multimodal, e.g. for readings recorded at a few discrete levels. A random forest [20], which makes no distributional assumption, answers the queries on which naive Bayes is uncertain:

$$q(c \mid \mathbf{x}) = \frac{1}{T} \sum_{t=1}^{T} \mathbb{1}\left[h_t(\mathbf{x}) = c\right]. \qquad (6)$$

M5: Confidence gate. The confidence of the primary stage is $s(\mathbf{x}) = \max_c P(c \mid \tilde{\mathbf{x}})$ (7), and the recommendation is

$$f_\gamma(\mathbf{x}) = \begin{cases} \underset{c}{\operatorname{argmax}} P(c \mid \tilde{\mathbf{x}}), & s(\mathbf{x}) \geq \gamma, \\ \underset{c}{\operatorname{argmax}} q(c \mid \mathbf{x}), & s(\mathbf{x}) < \gamma. \end{cases} \qquad (8)$$

The coverage $\varphi(\gamma) = \Pr[s(\mathbf{x}) \geq \gamma]$ (9) is the fraction of queries answered by naive Bayes, and the overall accuracy decomposes as

$$\mathrm{Acc}(\gamma) = \varphi(\gamma)\, A_G(\gamma) + \big(1 - \varphi(\gamma)\big) A_R(\gamma), \qquad (10)$$

where $A_G(\gamma)$ and $A_R(\gamma)$ are the accuracies of naive Bayes on accepted queries and of the forest on deferred queries. The threshold is chosen on a validation split held out from the training data,

$$\gamma^* = \min\left\{\gamma \in \Gamma:\ \mathrm{Acc}_{val}(\gamma) = \max_{\gamma' \in \Gamma} \mathrm{Acc}_{val}(\gamma')\right\}, \qquad (11)$$

over the grid $\Gamma = \{0, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95, 0.99, 0.999\}$; taking the smallest maximiser keeps as many queries as possible on the cheaper stage.

Proposition 1 (when deferral helps). Let $B(\gamma)$ be the accuracy of naive Bayes on the queries with $s(\mathbf{x}) < \gamma$. Since naive Bayes alone corresponds to $\gamma = 0$, Eq. (10) gives

$$\mathrm{Acc}(\gamma) - \mathrm{Acc}(0) = \big(1 - \varphi(\gamma)\big)[A_R(\gamma) - B(\gamma)]. \qquad (12)$$

The gate helps exactly when the forest is more accurate than naive Bayes on low-confidence queries; by Chow's analysis of the reject option [17], these are where the errors of a calibrated classifier concentrate.

Proposition 2 (reliability of the selected gate). Validation accuracy is an average of n_v bounded losses, so by Hoeffding's inequality and a union bound over $|\Gamma|$ candidates, with probability at least $1 - \delta$ every candidate is within

$$\varepsilon = \sqrt{\frac{\ln(2|\Gamma|/\delta)}{2n_v}} \qquad (13)$$

of its expected accuracy. For the Kaggle data ($n_v = 352$, $|\Gamma| = 9$, $\delta = 0.05$) $\varepsilon \approx 0.091$, and for the Maharashtra data ($n_v = 722$) $\varepsilon \approx 0.064$, so the selection resolves accuracy differences of several points; Section 4 therefore evaluates all methods over repeated runs.

The training and inference procedures are listed in Algorithms 1 and 2. In Algorithm 1, Lines 1–2 fit the scaler of Eq. (1), Lines 3–5 tune ϵ, Lines 6–8 fit both stages on the reduced training split, Lines 9–13 evaluate Eq. (10) for each gate and select γ^* by Eq. (11), and Lines 14–15 refit both stages on all training data.

Algorithm 1. Training of the confidence-gated recommender

Input: $\mathcal{D}_{tr}$; smoothing candidates E (20 draws); gate grid Γ; folds $K = 5$
Output: scaler, naive Bayes G, forest F, gate γ^*
1: store $\min_j, \max_j$ of every feature in $\mathcal{D}_{tr}$
2: transform $\mathcal{D}_{tr}$ by Eq. (1)
3: **for each** $\epsilon \in E$ **do** $a_\epsilon \leftarrow K$-fold accuracy of Eqs. (2)–(5)
4: **end for**
5: $\epsilon^* \leftarrow \operatorname{argmax}_\epsilon a_\epsilon$
6: split $\mathcal{D}_{tr}$ into $\mathcal{D}_a$ (80%) and $\mathcal{D}_{val}$ (20%), stratified
7: fit G on $\mathcal{D}_a$ with ϵ^* (Eq. 2); fit F on $\mathcal{D}_a$ (Eq. 6)
8: compute $s(\mathbf{x})$ (Eq. 7) for all $\mathbf{x} \in \mathcal{D}_{val}$
9: **for each** $\gamma \in \Gamma$ **do**
10: $\quad \hat{y} \leftarrow f_\gamma(\mathbf{x})$ on $\mathcal{D}_{val}$ (Eq. 8)
11: $\quad \mathrm{Acc}_{val}(\gamma) \leftarrow$ accuracy of $\hat{y}$ (Eq. 10)
12: **end for**
13: $\gamma^* \leftarrow$ smallest maximiser of Acc_{val} (Eq. 11)
14: refit G and F on all of $\mathcal{D}_{tr}$
15: **return** scaler, G, F, γ^*

Complexity of Algorithm 1. Eq. (2) costs $O(n\,d)$ time and $O(C\,d)$ memory, so the smoothing search costs $O(|E|\,K\,n\,d)$; the forest costs $O(T\,m\,n\log n)$ time and $O(T\,N_{node})$ memory; the gate search costs $O(|\Gamma|\,n_v)$.

In Algorithm 2, Lines 1–2 acquire and scale the reading, Lines 3–6 evaluate Eqs. (4)–(5) for every crop, Lines 7–14 apply the gate of Eq. (8), and Lines 15–16 return the recommendation with ranked alternatives.

Algorithm 2. Online crop recommendation

Input: reading **x** (form or IoT node); scaler; G; F; γ^*
Output: crop $\hat{y}$, confidence s, ranked list L
validate the ranges of N, P, K, T, H, pH, R
$\tilde{\mathbf{x}} \leftarrow$ scale **x** by Eq. (1)
for $c = 1$ to C **do**
  $g_c \leftarrow$ discriminant of Eq. (5)
end for
$P(c \mid \tilde{\mathbf{x}}) \leftarrow \exp(g_c) / \sum_{c'} \exp(g_{c'})$ (Eq. 4)
$s \leftarrow \max_c P(c \mid \tilde{\mathbf{x}})$ (Eq. 7)
if $s \geq \gamma^*$ **then**
  $\hat{y} \leftarrow \operatorname{argmax}_c P(c \mid \tilde{\mathbf{x}})$
  $L \leftarrow$ three crops with the largest P
else
  $\hat{y} \leftarrow \operatorname{argmax}_c q(c \mid \mathbf{x})$ (Eq. 6)
  $L \leftarrow$ three crops with the largest q
end if
display $\hat{y}$, s and L in the Streamlit interface
return $\hat{y}, s, L$

Complexity of Algorithm 2. The naive Bayes stage costs $O(C\,d)$ operations (154 multiply–add groups for the Kaggle data) and $O(C\,d)$ memory; only deferred queries incur the forest cost $O(T\,D)$ for average depth D.

3.3 Datasets, preprocessing and experimental protocol

Primary dataset. The Kaggle Crop Recommendation dataset [21] contains 2,200 records (100 per crop, 22 crops) with N (0–140), P (5–145), K (5–205), temperature (8.8–43.7 °C), humidity (14.3–100%), pH (3.5–9.9) and rainfall (20.2–298.6 mm), without missing values or duplicates.

Second dataset. The Crop and Fertilizer Dataset for Western Maharashtra [22] contains 4,513 records from five districts and 16 imbalanced crops (12 to 1,010 records). Nitrogen, phosphorus, potassium, pH, annual rainfall (300–1,700 mm) and temperature are recorded at few discrete levels (e.g. 7 distinct pH values); humidity is absent. Models are therefore re-tuned and retrained on it rather than transferred.

Protocol. Each dataset is split 80:20 by stratified sampling with ten seeds; results are mean ± standard deviation. All models are tuned once on the seed-0 training split with 20 random configurations and five-fold cross-validation (Table 2). Accuracy and the macro F1-score are reported,

$$\mathrm{F1}_{macro} = \frac{1}{C}\sum_{c=1}^{C} \frac{2\,\mathrm{Pre}_c\,\mathrm{Rec}_c}{\mathrm{Pre}_c + \mathrm{Rec}_c}. \qquad (14)$$

with two-sided Wilcoxon signed-rank and Friedman tests [23]. Permutation importance is the mean accuracy drop when feature j is permuted ($R = 30$),

$$I_j = \mathrm{Acc}(\mathcal{D}_{te}) - \frac{1}{R}\sum_{r=1}^{R} \mathrm{Acc}\left(\mathcal{D}_{te}^{(j,r)}\right). \qquad (15)$$

Table 2. Search spaces and selected hyperparameters (20 random configurations, five-fold cross-validation; Kaggle / Maharashtra).

Model	Search space	Kaggle	Maharashtra
LR	C ∈ [0.01, 1000], log scale	C = 287.7	C = 658.1
KNN	k = 1–15; uniform or distance weights	k = 6, distance	k = 1, uniform
DT	depth 8–20 or none; min. leaf 1–5; Gini/entropy	none, 1, entropy	none, 1, entropy
GNB	ε ∈ [1e-12, 0.1], log scale	ε = 1.1e-6	ε = 1.4e-3
RF (fallback)	T = 100–400; m ∈ {√d, $\log_2$d, 0.5d}; depth 15, 25, none; leaf 1–3	100, 0.5d, 25, 2	200, $\log_2$d, none, 1
SVM (RBF)	C ∈ [0.1, 1000], γ ∈ [0.01, 10], log scale	595.7, 1.11	821.2, 2.50
XGBoost	trees 100–300; depth 3–8; rate 0.02–0.3; subsample 0.8–1	200, 3, 0.192	200, 7, 0.104
Gate γ*	grid Γ (Section 3.2), 20% validation split	0 in 6 of 10 runs	0.95 in 8 of 10 runs

Baselines. Five recommenders from 2024–2026 were re-implemented: XGBoost (XGB-2024) [5]; a stacking ensemble built from the base models compared by Gunasekaran et al. (logistic regression, SVM, decision tree, naive Bayes, k-nearest neighbours, extra trees and XGBoost) with a logistic meta-learner (Stack-2024) [6]; the random forest of AgroXAI (RF-XAI-2024) [12]; the exponential-weighted ensemble of seven base models (ExpW-2025) [8], in which the weight of model k is $w_k \propto \exp(-\eta\, e_k)$ (16) for validation error e_k and η chosen from {1, 5, 10, 20, 50}; and the hybrid of random forest and SVM, implemented as soft voting (RF-SVM-2026) [14]. LR, KNN, DT and GNB complete the comparison; ensemble members reuse the tuned base settings.

Implementation. Python 3.13, scikit-learn 1.9.1 [24], XGBoost 3.4.1 and Streamlit on a two-core 2.10 GHz Intel Xeon CPU with 7 GB RAM; latency is the median of 300 single-thread queries.

4. Results and Discussion

4.1 Comparison with recent baselines

Table 3 and Figure 2 report the results on both datasets. On the Kaggle data the classical classifiers reach 98.20% (LR), 98.39% (KNN), 98.34% (DT) and 99.57% (GNB). The proposed recommender obtains the highest mean accuracy and macro F1-score (99.61 ± 0.30%), followed by RF-SVM-2026 (99.59%), GNB (99.57%) and ExpW-2025 (99.55%). The methods differ overall (Friedman $\chi_F^2 = 67.99$, $p < 10^{-10}$), but the proposed recommender, GNB, Stack-2024, ExpW-2025 and RF-SVM-2026 are statistically equivalent ($p \geq 0.25$), whereas LR, KNN, DT, XGB-2024 and RF-XAI-2024 are significantly less accurate ($p \leq 0.016$). The pooled confusion matrix and the confidence distributions of the gated recommender are shown in Figures 3 and 4.

Table 3. Accuracy and macro F1-score (mean ± standard deviation over ten runs, %). Best value in bold, second best underlined; p is the Wilcoxon signed-rank p-value of the accuracy difference with respect to the proposed recommender.

Model	Kaggle Acc.	Kaggle F1	p	Maharashtra Acc.	Maharashtra F1	p
LR	98.20 ± 0.42	98.20 ± 0.42	0.002	81.55 ± 1.05	72.89 ± 2.73	0.002
KNN	98.39 ± 0.71	98.38 ± 0.72	0.004	98.89 ± 0.31	98.10 ± 0.64	0.004
DT	98.34 ± 0.77	98.34 ± 0.77	0.002	97.59 ± 0.80	95.62 ± 1.31	0.002
GNB	99.57 ± 0.29	99.57 ± 0.29	0.500	79.05 ± 1.04	69.11 ± 1.89	0.002
XGB-2024 [5]	98.93 ± 0.22	98.93 ± 0.22	0.002	98.55 ± 0.41	97.56 ± 0.78	0.125
Stack-2024 [6]	99.50 ± 0.14	99.50 ± 0.14	0.250	**99.08 ± 0.23**	**98.50 ± 0.56**	0.002
RF-XAI-2024 [12]	99.41 ± 0.22	99.41 ± 0.22	0.016	98.37 ± 0.30	97.30 ± 0.62	1.000
ExpW-2025 [8]	99.55 ± 0.21	99.54 ± 0.21	0.438	98.96 ± 0.20	98.32 ± 0.47	0.002
RF-SVM-2026 [14]	99.59 ± 0.28	99.59 ± 0.28	0.469	98.63 ± 0.31	97.70 ± 0.57	0.008
Proposed	**99.61 ± 0.30**	**99.61 ± 0.30**	–	98.35 ± 0.32	97.29 ± 0.64	–

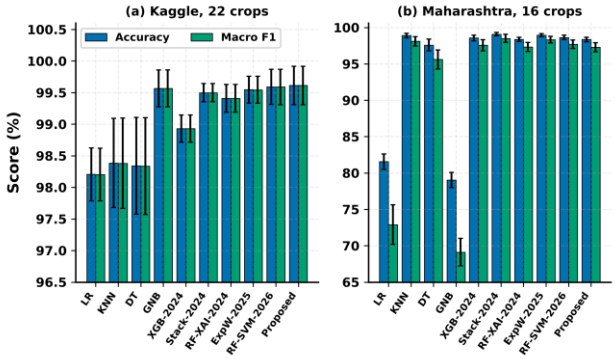


Figure 2. Accuracy and macro F1-score (mean ± standard deviation over ten runs) on (a) the Kaggle data and (b) the Maharashtra data. Note the different vertical scales: the ranking of naive Bayes changes from first to last between the two datasets, while the proposed recommender stays within one point of the best method on both.

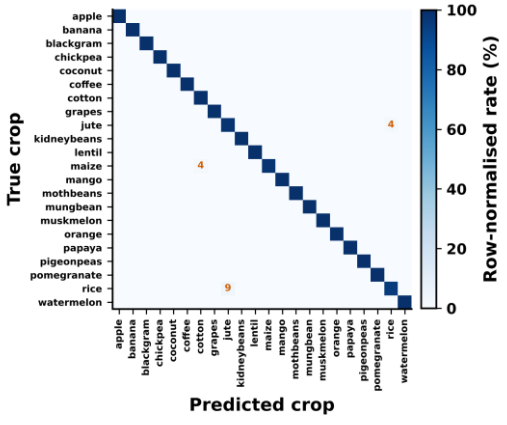


Figure 3. Confusion matrix of the proposed recommender on the Kaggle data, pooled over ten test splits. Off-diagonal counts are printed in red; the 17 errors involve mainly rice, jute, maize and cotton.

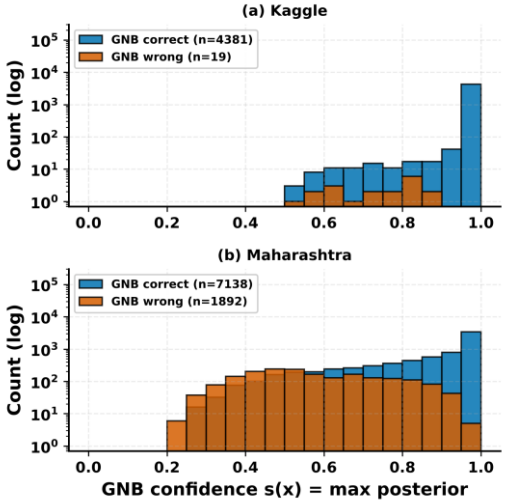


Figure 4. Distribution of naive Bayes confidence for correct (blue) and incorrect (orange) predictions, pooled over ten runs, on (a) Kaggle and (b) Maharashtra (log scale). Errors are rare and separable by confidence in (a) but frequent and spread across confidence levels in (b).

On the Maharashtra data naive Bayes falls to 79.05% (macro F1 69.11%) and LR to 81.55%, while tree and neighbour models stay above 97.5%. The proposed recommender recovers 19.30 points and reaches 98.35%, equal to the forest alone ($p = 1.0$);

Stack-2024 (99.08%), ExpW-2025 (98.96%), 1-nearest-neighbour (98.89%) and RF-SVM-2026 (98.63%) are significantly more accurate ($p \leq 0.008$); the result of 1-nearest-neighbour is consistent with the coarse readings, which leave many near-identical training records.

Figure 3 shows that only 17 of 4,400 pooled Kaggle predictions are wrong: rice as jute (9), jute as rice (4) and maize as cotton (4), crops with overlapping humidity and rainfall. Figure 4 explains the gate. On Kaggle, 4,381 of 4,400 naive Bayes predictions are correct, mostly with confidence above 0.95, so $\gamma^* = 0$ in six of ten runs (coverage 99.4%). On Maharashtra, 1,892 of 9,030 are wrong with confidences spread from 0.2 to 1.0, so $\gamma^* = 0.95$ and about 35% of queries remain with naive Bayes.

4.2 Ablation study

Table 4 and Figure 5 remove each module and input in turn. On Kaggle, removing the gate costs 0.04 points and removing the naive Bayes stage 0.20 points, while tuning of ϵ and scaling have no measurable effect. On Maharashtra, removing the gate costs 19.30 points and removing naive Bayes nothing, as Eq. (12) predicts when the forest is better on low-confidence queries. Rainfall is the most important input on both datasets (−3.18 and −8.78 points), followed by humidity on Kaggle (−1.31) and potassium and temperature on Maharashtra (−4.51, −4.27).

Table 4. Ablation results (accuracy, mean ± standard deviation over ten runs, %). Best value in bold, second best underlined.

Configuration	Kaggle	Maharashtra
Proposed (full)	**99.61 ± 0.30**	<u>98.35 ± 0.32</u>
w/o confidence gate (GNB only)	99.57 ± 0.29	79.05 ± 1.04
w/o GNB stage (RF only)	99.41 ± 0.22	**98.37 ± 0.30**
w/o tuning of ϵ	**99.61 ± 0.30**	<u>98.35 ± 0.32</u>
w/o min-max scaling	<u>99.59 ± 0.28</u>	**98.37 ± 0.30**
w/o nitrogen	99.34 ± 0.33	97.61 ± 0.53
w/o phosphorus	99.30 ± 0.36	94.84 ± 0.44
w/o potassium	99.11 ± 0.56	93.84 ± 0.96
w/o temperature	99.32 ± 0.28	94.08 ± 0.52
w/o humidity	98.30 ± 0.52	not available
w/o pH	99.34 ± 0.27	94.96 ± 0.73
w/o rainfall	96.43 ± 0.67	89.57 ± 1.13

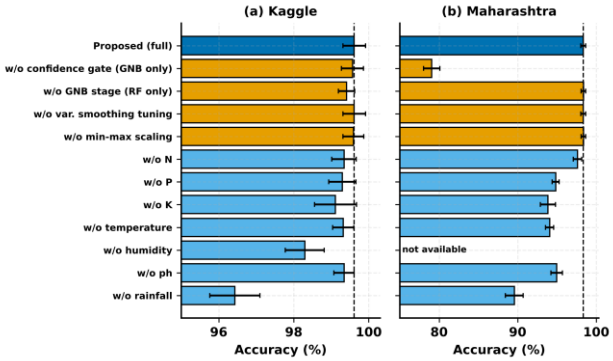


Figure 5. Ablation of modules (orange) and inputs (light blue) on (a) Kaggle and (b) Maharashtra; the dashed line marks the full recommender. Note the different horizontal scales. Humidity is not recorded in the Maharashtra data.

4.3 Sensitivity and convergence

Figure 6 shows that naive Bayes accuracy is constant for $\epsilon \leq 10^{-3}$ and drops only for $\epsilon \geq 10^{-2}$ (to 96.5–97.6% at 10^{-1}), whereas the training fraction matters more (98.4% at 10% to 99.7% at 80%). In Figure 7, naive Bayes reaches 99.31% with 20% of the data (about 18 records per crop), while DT and KNN need 60–80% to approach 98%, a consequence of the few parameters it estimates ($2Cd + C = 330$).

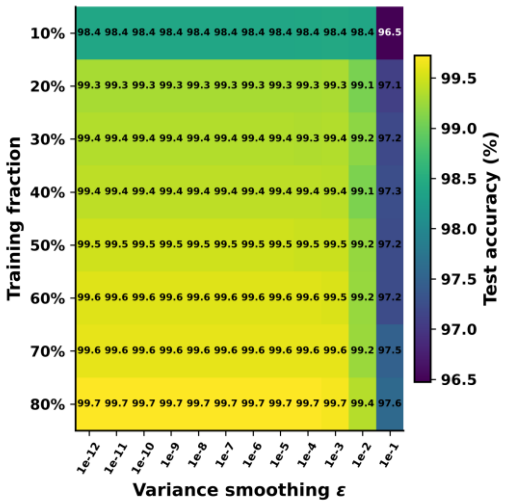


Figure 6. Test accuracy of naive Bayes on the Kaggle data versus variance smoothing ϵ and training fraction (mean of five splits). The model is insensitive to ϵ up to 10^{-3}.

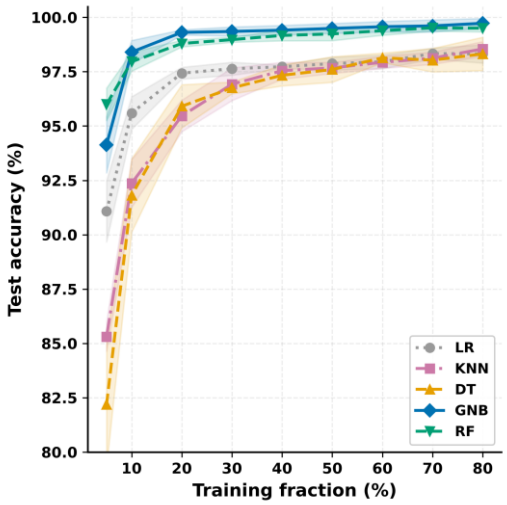


Figure 7. Learning curves (mean ± standard deviation over five splits) on the Kaggle data. Naive Bayes converges with about 20% of the training data, earlier than the other classifiers.

4.4 Interpretability

In Figure 8(a), permuting potassium (43.8 points), humidity (43.4) or rainfall (35.8) causes the largest accuracy losses, followed by phosphorus (18.1), nitrogen (15.6), temperature (9.1) and pH (5.4). The class means in Figure 8(b) explain why: potassium and phosphorus single out apple and grapes, humidity separates chickpea and kidney beans from coconut, papaya and the melons, and rainfall separates rice and coconut. The ranking differs from Table 4 because retraining lets correlated inputs compensate, but both analyses show that the water-related variables carry information the soil variables cannot replace, which is agronomically plausible [1], [10].

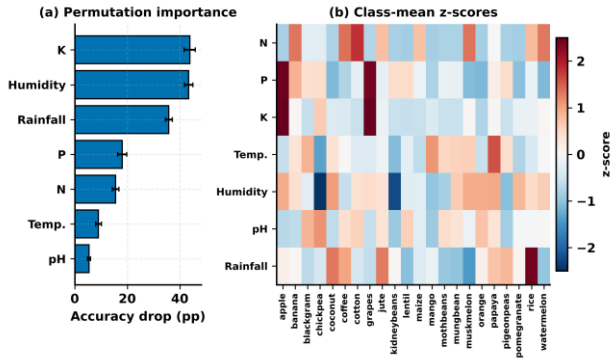


Figure 8. Interpretation of the naive Bayes stage on the Kaggle data. (a) Permutation importance (accuracy drop in percentage points, mean ± standard deviation over 30 permutations). (b) Class means of each input expressed as z-scores relative to the whole dataset; red cells mark crops that need unusually high values and blue cells unusually low values.

4.5 Computational cost and deployment

Table 5 lists training time, latency and model size. Naive Bayes trains in about 3 ms, answers in 0.3–0.4 ms and occupies 3–4 kB, whereas the stacking ensemble needs 63–91 s, about 18 ms per query and 50–80 MB. In Figure 9 the proposed recommender matches the most accurate ensembles at 0.92 ms per query (0.69 ms on average), 19 times faster than Stack-2024 and 9 times faster than ExpW-2025; on Maharashtra, where 65% of queries are deferred, its average latency is 7.5 ms.

Table 5. Computational cost on a single CPU thread (training time on 80% of the data, median latency of one query, serialised model size). Lowest value in bold, second lowest underlined.

Model	Kaggle train (s)	Kaggle latency (ms)	Kaggle size (kB)	Mah. train (s)	Mah. latency (ms)	Mah. size (kB)
LR	0.100	<u>0.22</u>	**3**	0.308	<u>0.22</u>	**2**
KNN	<u>0.003</u>	0.57	135	<u>0.004</u>	1.02	243
DT	0.010	**0.08**	21	0.005	**0.08**	92
GNB	**0.003**	0.40	<u>4</u>	**0.003**	0.33	<u>3</u>
XGB-2024 [5]	26.319	0.59	3,231	32.053	0.58	3,952
Stack-2024 [6]	62.890	17.80	50,041	91.105	17.57	79,592
RF-XAI-2024 [12]	0.274	4.86	2,479	0.574	9.53	26,267

Model	Kaggle train (s)	Kaggle latency (ms)	Kaggle size (kB)	Mah. train (s)	Mah. latency (ms)	Mah. size (kB)
ExpW-2025 [8]	31.498	8.09	5,998	55.464	12.27	30,703
RF-SVM-2026 [14]	0.435	5.75	2,605	1.294	10.37	26,411
Proposed	0.659	0.92	2,483	1.157	10.32	26,270

In the Streamlit application the user enters the seven readings, or a sensor node posts them as JSON, and receives the crop, its confidence and two alternatives. When $\gamma^* = 0$ the forest can be omitted, leaving a 4 kB model that can run offline on low-power IoT gateways [25]; otherwise the forest can reside on a phone or local server for deferred queries. Both stages run on a CPU in milliseconds, so hosting costs little or nothing.

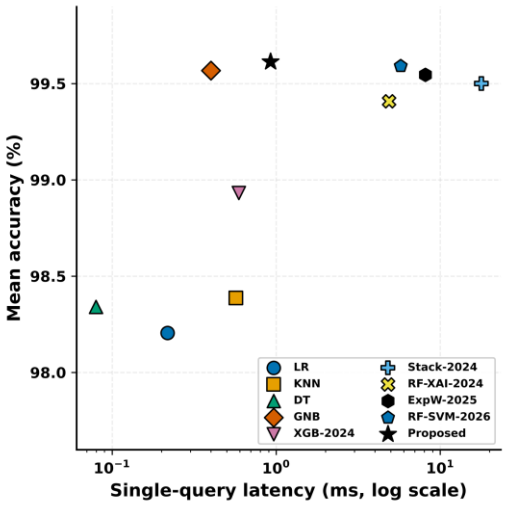


Figure 9. Mean accuracy versus single-query latency on the Kaggle data (one CPU thread, log scale). The proposed recommender (star) reaches the accuracy of the ensembles at about one millisecond per query.

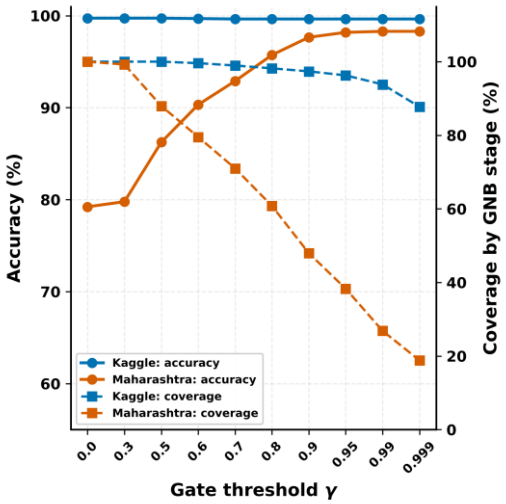


Figure 10. Accuracy (solid) and naive Bayes coverage (dashed) as functions of the gate threshold γ (mean of five splits). On Maharashtra, accuracy increases from 79.2% to 98.3% as more queries are deferred; on Kaggle it stays near 99.6% at all thresholds.

4.6 Failure-case analysis

Two failure types were observed. First, crops with overlapping climatic needs: all 17 Kaggle errors involve rice, jute, maize and cotton, and in every naive Bayes error the correct crop was ranked second (top-three accuracy 100%). Second, naive Bayes collapses on the Maharashtra data because discrete readings and several nutrient regimes per crop make the class-conditional distributions multimodal. Figure 10 shows the gate compensating, from 79.2% at $\gamma = 0$ to 98.3% at $\gamma \geq 0.95$, but with only 18–38% of queries left on the cheap stage; on these data the recommender operates mainly through the forest, and the stacking ensemble remains 0.73 points more accurate.

4.7 Limitations and threats to validity

Data characteristics: the Kaggle classes are balanced and well separated, so differences of a few tenths of a point among the leading methods lie within run-to-run variation. The Maharashtra data are imbalanced and coarsely recorded, and near-identical records on both sides of the split favour neighbour-based models there. Tuning and baselines: hyperparameters were tuned on the first split and then fixed for all methods, and the baselines follow their published descriptions. Gate selection: by Proposition 2 the validation split resolves differences of several points, which accounts for the variation of γ^* between Kaggle runs. Scope: market prices, irrigation and crop rotation are outside the model. Leakage was controlled by fitting the scaler, models and gate on training data only, and overfitting by cross-validated tuning and an untouched test split.

5. Conclusion

Under equal tuning, repeated runs and significance tests, a confidence-gated recommender combining Gaussian naive Bayes with a random-forest fallback was compared with four classical and five recent methods. On the Kaggle data, naive Bayes reached 99.57% accuracy and the gated recommender 99.61%, statistically equivalent to the best ensembles while answering a query in under one millisecond with a 4 kB model when the forest is not needed. On the Maharashtra data, naive Bayes fell to 79.05%, and the gate restored accuracy to 98.35%, an improvement of 19.30 points, although the stacking ensemble remained more accurate (99.08%). Rainfall, humidity and potassium were the most informative inputs.

Future work. Mixture or kernel densities will replace the single Gaussian per class, the gate will be calibrated on larger validation sets, and the system will be evaluated with IoT sensor readings over a full season.

Real-time applications. The recommender runs on IoT gateways, smartphones or advisory kiosks, returning a ranked crop list within milliseconds without network access, and can turn soil-health card results into immediate advice.

References

[1] K. G. Liakos, P. Busato, D. Moshou, S. Pearson and D. Bochtis, "Machine learning in agriculture: A review," *Sensors*, vol. 18, no. 8, p. 2674, 2018, doi: 10.3390/s18082674.

[2] A. Sharma, A. Jain, P. Gupta and V. Chowdary, "Machine learning applications for precision agriculture: A comprehensive review," *IEEE Access*, vol. 9, pp. 4843–4873, 2021, doi: 10.1109/ACCESS.2020.3048415.

[3] M. R. Islam, K. Oliullah, M. M. Kabir, M. Alom and M. F. Mridha, "Machine learning enabled IoT system for soil nutrients monitoring and crop recommendation," *Journal of Agriculture and Food Research*, vol. 14, p. 100880, 2023, doi: 10.1016/j.jafr.2023.100880.

[4] M. K. Senapaty, A. Ray and N. Padhy, "IoT-enabled soil nutrient analysis and crop recommendation model for precision agriculture," *Computers*, vol. 12, no. 3, p. 61, 2023, doi: 10.3390/computers12030061.

[5] B. Dey, J. Ferdous and R. Ahmed, "Machine learning based recommendation of agricultural and horticultural crop farming in India under the regime of NPK, soil pH and three climatic variables," *Heliyon*, vol. 10, no. 3, p. e25112, 2024, doi: 10.1016/j.heliyon.2024.e25112.

[6] H. Gunasekaran, D. Kanmani and R. Krishnamoorthi, "Optimized ensemble learning for enhanced crop recommendations: Leveraging ML for smarter agricultural decision-making," *Engineering Proceedings*, vol. 82, no. 1, p. 95, 2024, doi: 10.3390/ecsa-11-20366.

[7] Y. Akkem, S. K. Biswas and A. Varanasi, "Smart farming using artificial intelligence: A review," *Engineering Applications of Artificial Intelligence*, vol. 120, p. 105899, 2023, doi: 10.1016/j.engappai.2023.105899.

[8] L. Stracqualursi, "A lightweight exponential-weighted ensemble for crop recommendation," *Journal of Agricultural, Biological and Environmental Statistics*, 2025, doi: 10.1007/s13253-025-00694-6.

[9] T. van Klompenburg, A. Kassahun and C. Catal, "Crop yield prediction using machine learning: A systematic literature review," *Computers and Electronics in Agriculture*, vol. 177, p. 105709, 2020, doi: 10.1016/j.compag.2020.105709.

[10] T. A. Shaikh, T. Rasool and F. R. Lone, "Towards leveraging the role of machine learning and artificial intelligence in precision agriculture and smart farming," *Computers and Electronics in Agriculture*, vol. 198, p. 107119, 2022, doi: 10.1016/j.compag.2022.107119.

[11] E. Elbasi, C. Zaki, A. E. Topcu, W. Abdelbaki, A. I. Zreikat, E. Cina, A. Shdefat and L. Saker, "Crop prediction model using machine learning algorithms," *Applied Sciences*, vol. 13, no. 16, p. 9288, 2023, doi: 10.3390/app13169288.

[12] O. Turgut, I. Kok and S. Ozdemir, "AgroXAI: Explainable AI-driven crop recommendation system for Agriculture 4.0," arXiv preprint arXiv:2412.16196, 2024, doi: 10.48550/arXiv.2412.16196.

[13] R. A. Ahmed, W. El-Shafai, Z. A. Ahmed, E.-S. M. El-Rabaie and F. E. Abd El-Samie, "High-precision crop recommendation system with stacking ensemble classifiers for optimizing agricultural productivity," *Scientific Reports*, vol. 15, 2025, doi: 10.1038/s41598-025-09640-5.

[14] S. Dey, "A hybrid random forest and SVM framework for intelligent crop selection using soil health and environmental factors," *Science and Culture*, vol. 92, no. 3–4, pp. 268–269, 2026, doi: 10.36094/sc.v92.2026.A_Hybrid_Random_Forest_and_SVM_Framework.Dey.268.

[15] B. S. Rinayat and B. Dharaskar, "A robust ensemble approach for crop recommendation based on environmental parameters," *EPJ Web of Conferences*, vol. 328, p. 01024, 2025, doi: 10.1051/epjconf/202532801024.

[16] C. Musanase, A. Vodacek, D. Hanyurwimfura, A. Uwitonze and I. Kabandana, "Data-driven analysis and machine learning-based crop and fertilizer recommendation system for revolutionizing farming practices," *Agriculture*, vol. 13, no. 11, p. 2141, 2023, doi: 10.3390/agriculture13112141.

[17] C. K. Chow, "On optimum recognition error and reject tradeoff," *IEEE Transactions on Information Theory*, vol. 16, no. 1, pp. 41–46, 1970, doi: 10.1109/TIT.1970.1054406.

[18] Y. Geifman and R. El-Yaniv, "Selective classification for deep neural networks," in *Advances in Neural Information Processing Systems*, vol. 30, 2017.

[19] P. Domingos and M. Pazzani, "On the optimality of the simple Bayesian classifier under zero-one loss," *Machine Learning*, vol. 29, pp. 103–130, 1997, doi: 10.1023/A:1007413511361.

[20] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.

[21] A. Ingle, "Crop Recommendation Dataset," Kaggle, 2020. [Online]. Available: https://www.kaggle.com/datasets/atharvaingle/crop-recommendation-dataset

[22] "Crop and Fertilizer Dataset for Western Maharashtra," Kaggle (4,513 records, five districts, 16 crops). Copy used in this study: https://github.com/harishshivekar/AI-ML

[23] J. Demšar, "Statistical comparisons of classifiers over multiple data sets," *Journal of Machine Learning Research*, vol. 7, pp. 1–30, 2006.

[24] F. Pedregosa et al., "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[25] M. S. Farooq, S. Riaz, A. Abid, K. Abid and M. A. Naeem, "A survey on the role of IoT in agriculture for the implementation of smart farming," *IEEE Access*, vol. 7, pp. 156237–156271, 2019, doi: 10.1109/ACCESS.2019.2949703.

Siddhi Choudhary
CSE (Data Science) Department
GGSESTC, Bokaro

Shambhavi Mishra
CSE (Data Science) Department
GGSESTC, Bokaro

Suman Kumari
CSE (Cyber Security) Department
GGSESTC, Bokaro

Abstract

The rapid advancement of Artificial Intelligence (AI) has enabled autonomous AI agents to perform complex tasks, make decisions, interact with external tools, and adapt to dynamic environments with minimal human intervention. However, increasing autonomy raises concerns regarding transparency, reliability, behavioural consistency, and operational safety. This research explores AI Agent Observability and proposes a structured framework to monitor agent behaviour, analyse decision-making, track tool interactions, detect operational failures, and evaluate overall agent quality. The framework focuses on key parameters such as execution traces, response latency, task completion accuracy, tool-selection reliability, and error detection. By integrating continuous monitoring with systematic quality assessment, this study aims to make autonomous AI systems more transparent, reliable, accountable, and trustworthy.

Keywords:- AI agents, agentic systems, large language models, observability, evaluation benchmarks, AI safety, human oversight, guardrails

Introduction:

Autonomous Artificial Intelligence (AI) agents are rapidly transforming the way intelligent systems operate by enabling them to independently reason, plan, execute tasks, interact with external tools, and make decisions with minimal human intervention. However, as these systems become more autonomous, understanding and monitoring their behaviour becomes increasingly challenging. Unlike traditional AI models, an autonomous agent involves multiple interconnected operations, where a failure in reasoning, memory retrieval, tool selection, or task execution can affect the entire outcome. Evaluating only the final response is therefore insufficient to determine whether an AI agent has performed reliably.

The emergence of large language models (LLMs) with reasoning and tool-use capabilities has enabled a new class of systems commonly called AI agents. These systems can browse the web, write and execute code, manage files, call APIs, and coordinate with other agents to complete long-horizon tasks. As agents are given more responsibility, the question is no longer only "Is the answer correct?" but also "Did the agent take safe, efficient, and authorised actions to get there?"

Autonomy introduces three core difficulties:

Compounding errors. A small mistake in an early step can propagate through a long chain of actions.

Real-world side effects. Agents can send emails, modify databases, or spend money, so errors may be costly or irreversible.

Opacity. The internal reasoning of an agent is often hard to inspect, making it difficult to know why a particular action was taken.

AI Agent Observability addresses these challenges by providing continuous insights into an agent's operations, interactions, decisions, and failures. While conventional monitoring focuses primarily on system performance, observability aims to understand not only what happened, but also how and why it happened.

This paper investigates three research questions:

How can the internal behaviour and execution processes of autonomous AI agents be effectively monitored?

Which evaluation parameters can accurately measure agent reliability, performance, and behavioural consistency?

How can continuous observability and quality assurance help identify, analyse, and prevent failures in autonomous AI systems?

The contribution of this research is a proposed integrated framework for AI agent observability, behavioural analysis, and quality assurance that connects operational monitoring with systematic performance evaluation to support more transparent, reliable, and accountable autonomous AI systems.

Methodology

The proposed research adopts a structured, framework-based approach to examine the observability, behaviour, and reliability of autonomous AI agents. It focuses on monitoring their execution processes, analysing operational activities, identifying failures, and evaluating overall performance. The methodology consists of the following stages:

Agent Execution Monitoring: Monitor the complete execution lifecycle of autonomous AI agents, including task initiation, planning, reasoning, decision-making, and final task completion. This helps understand how agents perform operations and respond to different situations.

Execution Tracing and Logging: Maintain detailed records of agent activities, including API requests, tool interactions, information retrieval, execution sequences, response generation, and system errors. These records provide traceability and help reconstruct the sequence of events during agent execution.

Behavioural Analysis: Examine agent behaviour to identify decision-making patterns, repeated actions, incorrect tool selection, unexpected responses, and unnecessary execution loops. This stage helps detect behavioural inconsistencies that may affect the agent's reliability.

Performance Evaluation: Evaluate agent performance using measurable parameters such as task completion rate, response latency, execution time, tool-selection accuracy, and error frequency. These metrics help assess the efficiency and consistency of autonomous operations.

Failure Detection and Analysis: Identify and analyse failures such as unsuccessful tool calls, incorrect information retrieval, reasoning errors, and incomplete task execution. Understanding the possible causes of these failures helps determine their impact on the overall functioning of an AI agent.

Continuous Quality Assurance: Integrate monitoring results, behavioural observations, and performance metrics into a continuous evaluation process. The collected insights can support corrective actions, improve operational strategies, and strengthen the reliability, transparency, and accountability of autonomous AI systems.

The proposed methodology brings together operational monitoring, behavioural analysis, performance measurement, and systematic failure assessment to establish a comprehensive approach to AI Agent Observability.

Literature Review

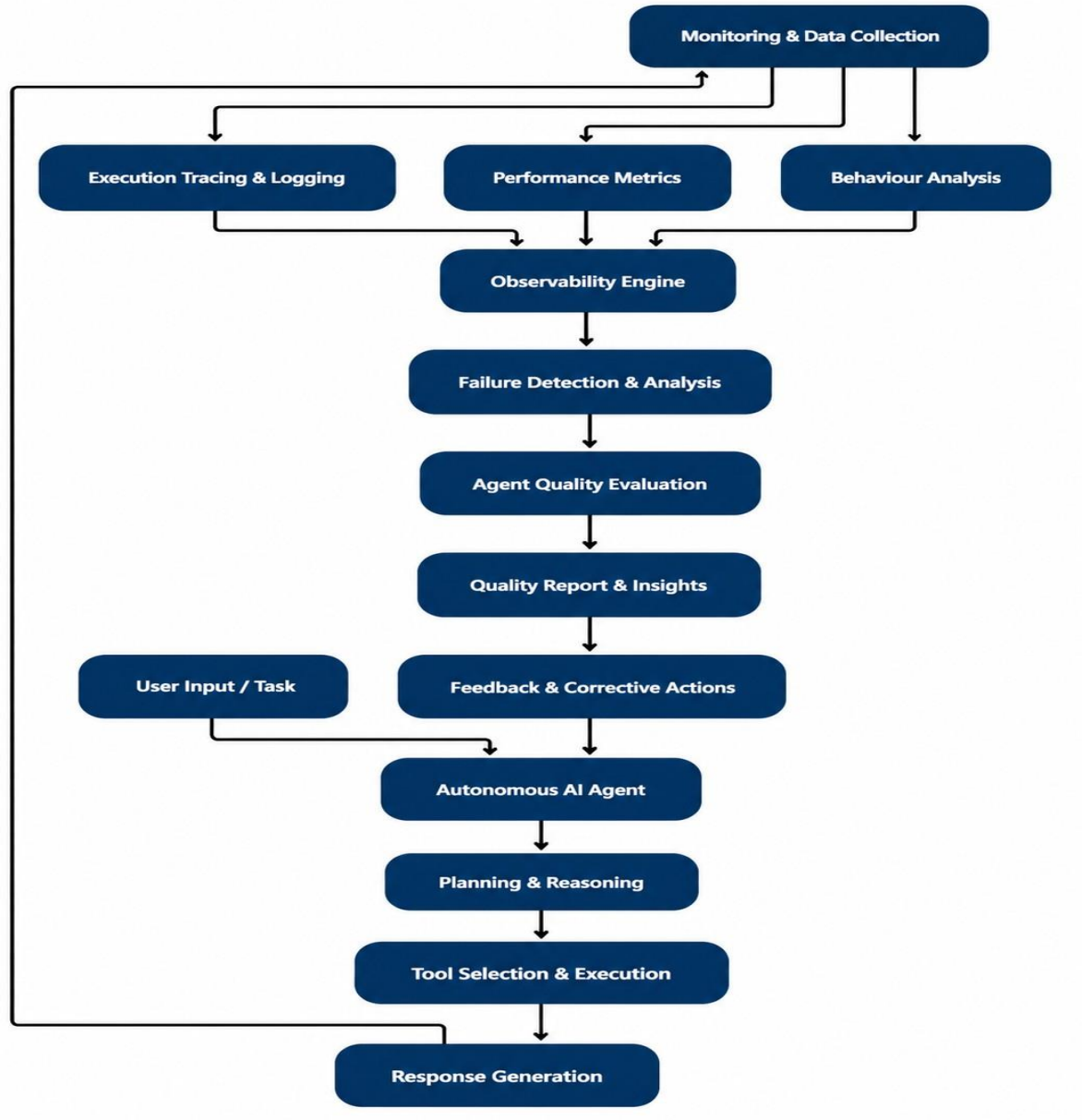
Recent research on autonomous AI agents has increasingly focused on their evaluation, reliability, and transparency. Yehudai et al. (2025) [1] examined agent evaluation methods, highlighting challenges in assessing planning, tool usage, memory, and overall performance. Similarly, Mohammadi et al. (2025) [2] explored different approaches to evaluating agent behaviour, capabilities, reliability, and safety. Benchmarks such as AgentBench [6] evaluate LLMs acting as agents across interactive environments. Recent work on evidence tracing and execution provenance further emphasises the importance of tracking agent actions, retrieved information, and decision processes to identify the causes of failures.

Research on agent observability also highlights the limitations of conventional monitoring in detecting failures across complex agent operations [3], [4], [5]. Despite these developments, challenges remain in integrating behavioural monitoring, failure detection, execution tracing, and quality evaluation into a unified framework.

This research addresses these limitations by proposing an integrated approach to AI agent observability and systematic quality assurance.

System Design

The proposed architecture illustrates how an autonomous AI agent can be continuously monitored and evaluated throughout its execution. Initially, the agent receives a task, performs planning and reasoning, selects appropriate tools, and generates a response. During this process, an observability layer collects execution logs, performance metrics, and behavioural information. These records are analysed to identify operational failures, unexpected behaviour, and performance issues. The agent is then evaluated using quality parameters, and the resulting reports provide insights for corrective actions and continuous improvement. This creates a feedback loop that helps improve the reliability, transparency, and accountability of



autonomous AI systems.

Implementation Detail

The proposed AI Agent Observability Framework consists of six major components:

Agent Integration: Connects the autonomous AI agent with the monitoring framework to track its activities, decisions, and task execution.

Data Collection: Collects important information such as user inputs, agent responses, tool interactions, execution time, and errors.

Logging and Tracing: Records each execution step in sequence, helping understand the agent's workflow and identify where failures occur.

Behaviour Analysis: Examines agent behaviour to detect repeated actions, incorrect tool selection, unexpected responses, and unnecessary execution loops.

Performance Evaluation: Measures agent performance using parameters such as task completion rate, response time, accuracy, and error frequency.

Failure Detection and Reporting: Identifies operational problems, analyses their possible causes, and generates reports highlighting areas that require improvement.

Conclusion

Autonomous AI agents are becoming increasingly capable, making their monitoring and evaluation essential. This research proposes an AI Agent Observability Framework to track agent behaviour, evaluate performance, and detect failures. By combining continuous monitoring with quality assessment, the proposed approach can help make AI agents more reliable, transparent, and accountable.

References

[1] A. Yehudai et al., "Survey on Evaluation of LLM-based Agents," arXiv:2503.16416, 2025.

[2] M. Mohammadi, Y. Li, J. Lo, and W. Yip, "Evaluation and Benchmarking of LLM Agents: A Survey," 2025.

[3] D. Moshkovich et al., "Beyond Black-Box Benchmarking: Observability, Analytics, and Optimization of Agentic Systems," 2025.

[4] L. Dong, Q. Lu, and L. Zhu, "AgentOps: Enabling Observability of LLM Agents," 2024.

[5] K. C. Balusu, "AgentTelemetry: A Fault Detection Benchmark and Toolkit for LLM Agent Observability," in Proc. 3rd ACM International Conference on AI-Powered Software, 2026, pp. 380–387.

[6] X. Liu et al., "AgentBench: Evaluating LLMs as Agents," in Proc. International Conference on Learning Representations (ICLR), 2024.

AI-based digital payment fraud detection and prevention system

Lilawati Kumari
Government Women's Polytechnic, Bokaro, Jharkhand

Kumkum Kumari
Government Women's Polytechnic, Bokaro, Jharkhand

Neha Kumari
Government Women's Polytechnic, Bokaro, Jharkhand

ABSTRACT

This paper presents a Python-based educational prototype for strengthening digital-payment account security through layered authentication and rule-based risk analysis. The implemented prototype includes password hashing, device and IP-address checks, risk scoring using predefined rules, email-based one-time password (OTP) verification, account locking after repeated failed attempts, and transaction monitoring. It also demonstrates basic banking-style operations such as balance inquiry, money transfer, transaction-PIN validation, and transaction-history viewing. The current implementation does **not** train or use a machine-learning model and is not connected to a live bank, UPI service, or payment gateway. A proposed future enhancement is to add transaction-anomaly detection using features such as amount, time, frequency, and deviation from a user's historical pattern. Any flagged activity should trigger verification or review rather than be treated as proof of fraud.

Keywords: Digital payment security, fraud detection, risk-based authentication, Python, OTP, transaction monitoring, anomaly detection, machine
learning.

1. INTRODUCTION

Digital payments make transfers and everyday purchases fast and accessible, but digital financial services must also address risks such as stolen credentials, unauthorised account access, suspicious login attempts, and unusual transaction behaviour. A login process based only on a username and password may provide limited protection if credentials are exposed. A stronger approach can evaluate multiple signals and request additional verification when activity appears unusual.

The proposed project explores a layered approach to account security. Its current Python prototype checks user credentials, compares device and IP information with previously stored details, calculates a rule-based risk score, and requests OTP verification when configured conditions indicate elevated risk. It also records login and transaction activity.

2. PROBLEM STATEMENT AND OBJECTIVES

A change in device or IP address can be legitimate—for example, when a user changes a phone or network—so it should be treated as a risk indicator, not conclusive evidence of fraud. The project aims to combine several checks and apply proportionate verification.

PROJECT OBJECTIVES	EXPECTED SYSTEM ROLE
Protect user registration and login	Hash passwords before storing credentials
Identify unusual login context	Compare device and IP information with stored details
Apply risk-based verification	Calculate a rule-based score and request OTP when required
Monitor account activity	Record login events and demonstrate transaction-history functions

3. SYSTEM ARCHITECTURE AND WORKING

The system combines authentication checks, context-based risk assessment, OTP verification, account controls, and transaction monitoring. The current implementation is rule-based; the machine-learning module shown below is a proposed

3.1 Registration and credential verification

During registration, the prototype hashes the entered password using SHA-256 and stores the resulting value in a JSON file. At login, the entered password is hashed and compared with the stored hash. SHA-256 demonstrates the concept of hashing, but it is not recommended as a production password-storage method because it is a fast general-purpose hash. A production system should use a dedicated password-hashing method such as Argon2id or bcrypt, configured with appropriate salts.

3.2 Device and IP-address checks

The prototype compares current device or operating-system information and the current IP address with previously stored values. A newly observed device or changed IP can increase the risk score. These indicators are imperfect and may generate false alerts; therefore, they should prompt additional checks rather than automatically classify a user or transaction as fraudulent.

3.3 Rule-based risk score and OTP verification

The detection module evaluates selected conditions and calculates a risk score. Depending on the configured rules, the application may allow a normal login, request an email OTP, or restrict access. OTP adds a verification layer. A secure deployment would also require OTP expiry, attempt limits, protected delivery, and careful handling of secrets.

3.4 Account controls and transaction monitoring

Repeated unsuccessful login attempts can lead to account locking until administrator intervention. The banking module demonstrates balance inquiry, money transfer, transaction-PIN validation, transaction-history viewing, and temporary restrictions after repeated incorrect PIN attempts. These are prototype demonstrations only; they do not communicate with a real bank, UPI network, or payment gateway.

4. PROPOSED AI/ML ENHANCEMENT

To extend the rule-based prototype toward AI-assisted fraud detection, a separate transaction-anomaly module can analyse transaction amount, time, frequency, device context, and deviation from the account's historical pattern. One possible experimental approach is Isolation Forest, an unsupervised anomaly-detection algorithm. It can produce an anomaly score for review. The score should not be presented as a final fraud verdict.

Proposed workflow: Transaction input →Feature preparation →ML anomaly analysis →Combined risk assessment →Alert / additional verification / manual review

A labelled or otherwise suitable dataset, documented preprocessing, and careful evaluation would be required before claiming fraud-detection accuracy. This ML enhancement is proposed future work, not a feature of the current prototype.

5. TECHNOLOGIES AND SYSTEM MODULES

COMPONENT	ROLE IN THE PROTOTYPE
Python	Core implementation language
JSON files	Prototype data storage for user/account records
hashlib	Password hashing demonstration
platform / socket	Device and network-context information
smtplib / email.message	Email notifications and OTP delivery
datetime / random	Timestamps and prototype OTP generation

6. DEMONSTRATION AND EVALUATION PLAN

The prototype can be demonstrated with controlled test accounts and simulated scenarios. The table below describes expected behaviour to verify during testing; it does not report measured accuracy or completed independent evaluation.

TEST SCENARIO	SYSTEM CHECK	EXPECTED RESPONSE
Known device and familiar IP	Credential and context checks	Normal access, subject to configured rules
New device	Device mismatch	Risk score increases; OTP may be requested
Changed IP address	IP mismatch	Risk score increases; additional check may be
Repeated incorrect passwords	Failed-attempt counter	Account locks at configured limit
Incorrect OTP	OTP comparison	Login is denied

For a stronger evaluation, future testing should record the number of test cases, successful detections, false alerts, missed suspicious cases, and response time. If a labelled dataset is used, suitable measures may include precision, recall, F1-score, and a confusion matrix.

7. LIMITATIONS, PRIVACY AND FUTURE SCOPE

The current version is an educational prototype. JSON storage is not suitable for concurrent, high-volume financial systems.
Device and IP changes can produce false alerts, and rule-based scores depend on the conditions selected by the developers.
The system has not been validated on a large, representative, labelled digital-payment dataset; therefore, no real-world
fraud-detection accuracy or reduction in financial loss is claimed.

FUTURE IMPROVEMENT	PURPOSEm
Add and evaluate an ML anomaly model	Explore unusual transaction patterns using suitable data
Use secure database and stronger password hashing	Improve storage and credential protection
Add OTP expiry, rate limits and secure secret handling	Strengthen verification controls
Add admin review, audit logs and privacy safeguards	Support accountable monitoring and incident review

8. CONCLUSION

The project demonstrates a layered-authentication prototype that combines password verification, device and IP checks, rule-based risk scoring, OTP verification, account locking, and transaction-history features. These functions provide a practical foundation for identifying potentially suspicious access conditions. The current implementation is not a complete AI-powered banking solution; adding and evaluating a machine-learning anomaly-detection module is a future step. Clear separation between implemented features and proposed enhancements makes the project technically transparent and suitable for further development.

REFERENCES

1.OWASP Foundation. *Password Storage Cheat Sheet*.
https://cheatsheetseries.owasp.org/cheatsheets/Password_Storage_Cheat_Sheet.html

2. scikit-learn documentation. *IsolationForest*. https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.IsolationForest.html

3. Python Software Foundation. *Python Documentation*. https://docs.python.org/3/

Artificial Intelligence and Security: Threats, Defenses, and Governance

Sourabh Kumar[1], Ayush Singh[2], Md Hussain Ansari[3]

[1]Student, Department of CSE, Guru Gobind Singh Educational Society's Technical Campus, Bokaro
[2]Student, Department of CSE, Guru Gobind Singh Educational Society's Technical Campus, Kandra, Bokaro
[3]Assistant Professor, Department of CSE, Guru Gobind Singh Educational Society's Technical Campus, Kandra, Bokaro,
*Email of the corresponding author: (**Md Hussain Ansari**) **mdhussain.ansari@ggsestc.ac.in**

Abstract

Artificial intelligence (AI) now sits on both sides of the security boundary: it is a high-value asset that must be defended, and a capability that attackers and defenders both use. This paper examines that two-sided relationship through a structured review of the adversarial machine learning literature, standards bodies' guidance, and documented attacks on large language models (LLMs). We organize the threat landscape by lifecycle stage and attacker goal, describe how AI strengthens detection and response while also lowering the cost of offensive operations, and propose a layered defense-and-governance framework that combines secure development, runtime controls, and organizational oversight. We argue that no single technical defense is sufficient, that LLM-based agents widen the attack surface in ways classical software security does not anticipate, and that security evidence should be gathered continuously rather than at release time only. The paper closes with limitations and a research agenda.

Keywords: artificial intelligence; cybersecurity; adversarial machine learning; data poisoning; prompt injection; large language models; AI governance; risk management

1. Introduction

AI systems are now embedded in fraud screening, medical triage, vehicle perception, software development and customer service, so a failure of their security is a failure of the services built on them. Security research on AI is commonly split into two questions. The first asks how AI systems themselves can be attacked and protected. The second asks how AI changes the work of attackers and defenders in conventional computing. Treating them separately hides the fact that they interact: a defensive model that can be evaded gives false assurance, and a code-writing assistant that can be manipulated becomes an entry point into the very systems it serves.

This paper makes three contributions. First, it presents a lifecycle-based view of attacks on AI systems that covers both classical predictive models and generative models. Second, it summarizes the dual-use character of AI in cyber operations and identifies where defensive gains are most credible. Third, it proposes a layered framework that connects technical controls to governance practice, drawing on publicly available risk-management guidance [1], [2].

The paper is organized as follows. Section 2 gives background and related work. Section 3 states the review method. Section 4 analyzes threats to AI systems. Section 5 covers AI as a security tool. Section 6 presents the defense framework. Section 7 discusses limitations, and Section 8 concludes.

2. Background and Related Work

The vulnerability of learned models to small, deliberately chosen input changes was first reported for image classifiers, where imperceptible perturbations caused confident misclassification [3]. Subsequent work explained the effect through the near-linear behavior of high-dimensional models and introduced fast gradient-based methods for generating such inputs [4]. Adversarial training, in which a model is optimized against worst-case perturbations, became the most widely studied empirical defense [5]. Even so, defenses have repeatedly been weakened by stronger or adaptive attacks, which is why evaluation methodology matters as much as the defense itself.

A second line of work targets training rather than inference. Backdoor or trojan attacks implant a trigger so that the model behaves normally on clean inputs but follows an attacker's choice when the trigger appears [6]. More recent evidence suggests that poisoning web-scale datasets is practical when training data is gathered from sources an adversary can edit [15]. A third line concerns confidentiality: membership inference attacks reveal whether a record was used in training [7], and model-extraction attacks approximate a proprietary model through its prediction interface [8].

For generative models, privacy leakage has been shown directly. Researchers recovered verbatim training sequences, including personal information, from a deployed language model, and found that larger models were more exposed [9]. A later study showed that an aligned production chatbot could be induced to emit memorized data, indicating that alignment does not remove memorization [10]. Applications built around LLMs add a distinct risk: instructions hidden in retrieved content can steer the model, an attack known as indirect prompt injection [11].

On the standards side, NIST has published a taxonomy and terminology for adversarial machine learning, covering predictive and generative systems and organizing attacks by system type, lifecycle stage, attacker goal, capability and knowledge [1]. The guidance states plainly that no foolproof defense exists, and that existing mitigations are incomplete. NIST's AI Risk Management Framework supplies the governance vocabulary of govern, map, measure and manage [2]. The OWASP project maintains a community-ranked list of ten critical risks for LLM applications, in which prompt injection ranks first and categories such as supply chain, data and model poisoning, and excessive agency appear alongside it [12]. On the defensive-use side, a well-known critique argued that anomaly detection based on machine learning performs poorly in network intrusion settings because of the open-world nature of traffic and the high cost of errors [13], and a multi-institution report anticipated how AI would expand malicious capabilities [14].

3. Methodology

This is a qualitative, structured literature review with a conceptual synthesis; it reports no new experiments. Sources were drawn from three groups: peer-reviewed or conference papers on adversarial machine learning and LLM security, standards and guidance documents from public bodies and community projects, and position papers on malicious use of AI. Priority was given to primary sources, and to works that are widely cited or that define a taxonomy used by practitioners. Trade-press and vendor summaries were used only to locate primary documents, not as evidence.

The synthesis follows the dimensions used in the NIST taxonomy: the lifecycle stage at which an attack occurs, the property it violates (availability, integrity or privacy), and the knowledge and access the attacker is assumed to have [1]. Each attack family in Section 4 is described along these dimensions, and each control in Section 6 is mapped to the stage it protects.

Two limits of this design should be noted. A literature review reflects what has been published, which favors well-studied image and text models over emerging modalities. It also cannot measure real-world incident frequency, because most incidents are not publicly disclosed.

4. Threats to AI Systems

Attacks on AI systems fall into four families that differ by when they act and what they violate. Table 1 summarizes them; the paragraphs below explain the points that matter most for practice.

Attack family	Lifecycle stage	Property violated	Typical attacker access	Representative work
Evasion (adversarial examples)	Deployment	Integrity	Query or white-box access	[3], [4], [5]
Poisoning and backdoors	Data collection,	Integrity,	Ability to influence	[6], [15]

Attack family	Lifecycle stage	Property violated	Typical attacker access	Representative work
	training	availability	training data or supply chain	
Privacy attacks (membership inference, extraction, model theft)	Deployment	Privacy, confidentiality	Query access	[7], [8], [9], [10]
LLM application attacks (prompt injection, misuse)	Deployment, integration	Integrity, privacy, safety	Control of any text the model reads	[11], [12]

Table 1. Attack families on AI systems, organized by the dimensions used in the NIST taxonomy [1].

4.1 Evasion

Evasion changes inputs at inference time so that a model decides incorrectly while a human sees nothing unusual [3], [4]. The risk is greatest where the input channel is open to the adversary, such as uploaded images, documents, audio, or network traffic. Adversarial training improves robustness against the attack types it was trained on, but robustness rarely transfers to new perturbation types, and it often costs clean accuracy [5]. Practitioners should therefore treat robustness claims as scoped to a stated threat model, not as a general property.

4.2 Poisoning and backdoors

Poisoning moves the attack earlier, into the data or the pipeline. A backdoored model can pass ordinary validation because its flaw activates only on an attacker-chosen trigger [6]. The practical exposure grows with reliance on scraped data, third-party pre-trained models and shared datasets, since the provenance of each is difficult to verify; evidence indicates that adversaries can realistically alter portions of web-scale corpora [15]. This explains why OWASP lists both supply chain and data and model poisoning as separate top risks for LLM applications [12].

4.3 Privacy and extraction

Models can reveal information about their training sets. Membership inference tests whether a given record was used in training [7], and extraction attacks recover verbatim passages from language models, with larger models more exposed [9]. Work on production chatbots shows that alignment training does not by itself prevent leakage of memorized data [10]. Separately, an attacker who can query a prediction service may reconstruct a close copy of the model, undermining intellectual property and enabling stronger follow-on evasion attacks [8].

4.4 LLM applications and agents

LLM-based systems blur the boundary between instructions and data, because both arrive as natural language. In indirect prompt injection, an attacker plants instructions in a web page, email or document that the application later retrieves, and the model may follow them as if the user had written them [11]. The consequences depend on what the application is permitted to do. A model that only answers questions can leak information; a model with tools for sending messages, executing code or modifying records can cause real-world harm. OWASP's guidance reflects this by ranking prompt injection first and by expanding its treatment of excessive agency as agentic designs spread [12]. The central design lesson is that model output should be handled as untrusted input by every downstream component.

5. AI as a Security Tool and Dual-Use Risks

5.1 Defensive uses

Machine learning supports defenders in tasks where data volume exceeds human capacity: malware classification, phishing detection, log triage, user-behavior analytics and vulnerability prioritization. The strongest results come from supervised settings with stable labels, such as known malware families or spam. The weakest come from open-ended anomaly detection on live networks, where normal behavior shifts constantly and a rare false alarm already

costs analyst time [13]. A defensible position is that AI should rank and enrich alerts for human analysts instead of acting alone on high-impact decisions.

5.2 Offensive and dual-use uses

The same capabilities lower the cost of attack. Early assessments warned that AI would expand the set of people able to carry out attacks, make existing attacks cheaper and more targeted, and introduce new attack types [14]. In practice, the clearest effects are on scale and persuasion: tailored phishing text, automated reconnaissance, and faster adaptation of known techniques. Fluent generative models also reduce the language and skill barriers that once limited some campaigns.

5.3 A feedback loop

The two roles are linked. A detection model trained on yesterday's attacks can be evaded by an adversary who probes it, which places the defender's classifier inside the threat model of Section 4. Conversely, defenders who deploy LLM assistants for triage or code review inherit prompt-injection and data-leakage risk. Security teams should therefore apply the same adversarial testing to their own AI tools that they would require of any other exposed system.

6. Defense and Governance Framework

Because individual defenses are incomplete [1], we propose layering controls so that a failure at one layer is caught at another. The framework has four layers, each mapped to the lifecycle stage it protects and to the NIST govern, map, measure and manage functions [2].

Layer	Lifecycle stage	Representative controls	RMF function
1. Data and supply chain	Collection, training	Provenance records for datasets and models; integrity checks; vetting of third-party components; filtering and anomaly review of training data	Map, Manage
2. Model hardening	Training, evaluation	Adversarial training scoped to a stated threat model; backdoor scanning; privacy-preserving training; limits on memorization	Measure, Manage
3. Runtime and application controls	Deployment	Input validation; output filtering and redaction; rate limiting on queries; least-privilege access for tools; human approval for high-impact actions	Manage
4. Governance and assurance	Whole lifecycle	Named risk owners; threat modeling; red-teaming before and after release; incident response; logging and continuous monitoring	Govern, Measure

Table 2. Layered defense and governance framework.

Three principles cut across the layers. First, **assume the model can be manipulated**: treat its outputs as untrusted and restrict what it is able to do, in line with OWASP's emphasis on limiting tools, functions and permissions [12]. Second, **state the threat model**: a robustness or privacy claim is meaningful only relative to a defined attacker with defined access [1]. Third, **measure continuously**: test results at release decay as attackers adapt, so red-team findings and monitoring data should feed back into design, and incidents should be reviewed in the same way as incidents in conventional software.

7. Discussion and Limitations

The review points to a gap between what is demonstrated in research and what is measured in production. Most attack papers assume a defined model and dataset; deployed systems combine several models, retrieval stores, tools and human workflows, so the real attack surface is the whole pipeline. This favors system-level threat modeling over model-level robustness scores. The framework in Section 6 is a synthesis, not a validated method. It has not been tested in a controlled study, and the mapping of controls to lifecycle stages reflects judgment about the cited sources. The review also has the limits stated in Section 3: it depends on published work, undercounts undisclosed incidents, and moves quickly, since LLM and agent security guidance is revised frequently. Readers should check current editions of standards before relying on them.

Trade-offs deserve more attention than they receive. Stronger robustness can reduce accuracy, privacy-preserving training can reduce utility, and tight output filtering can frustrate legitimate users. A governance program should record which trade-offs were accepted, by whom, and on what evidence.

Future research. Priorities include benchmarks for agentic systems that capture multi-step tool use, methods for verifying data provenance at web scale, measurable criteria for assessing memorization risk, and field studies of how security teams actually use AI assistants.

8. Conclusion

AI security is best understood as a problem of both protecting AI and using AI, and the two cannot be separated in practice. Attacks on models span the whole lifecycle, from poisoned data to manipulated prompts, and published defenses remain incomplete. LLM-based agents raise the stakes because they combine untrusted text with the authority to act. We recommend a layered approach that limits what models can do, states threat models explicitly, and treats testing and monitoring as continuing duties. Future work should validate the proposed framework empirically and track how guidance evolves as systems become more autonomous.

References

1. National Institute of Standards and Technology, *Adversarial Machine Learning: A Taxonomy and Terminology of Attacks and Mitigations*, NIST AI 100-2e2025, Gaithersburg, MD, 2025. doi:10.6028/NIST.AI.100-2e2025
2. National Institute of Standards and Technology, *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, NIST AI 100-1, 2023.
3. C. Szegedy et al., "Intriguing properties of neural networks," in *Proc. Int. Conf. Learning Representations (ICLR)*, 2014.
4. I. J. Goodfellow, J. Shlens, and C. Szegedy, "Explaining and harnessing adversarial examples," in *Proc. ICLR*, 2015.
5. A. Madry, A. Makelov, L. Schmidt, D. Tsipras, and A. Vladu, "Towards deep learning models resistant to adversarial attacks," in *Proc. ICLR*, 2018.
6. T. Gu, B. Dolan-Gavitt, and S. Garg, "BadNets: Identifying vulnerabilities in the machine learning model supply chain," arXiv:1708.06733, 2017.
7. R. Shokri, M. Stronati, C. Song, and V. Shmatikov, "Membership inference attacks against machine learning models," in *Proc. IEEE Symp. Security and Privacy*, 2017.
8. F. Tramèr, F. Zhang, A. Juels, M. K. Reiter, and T. Ristenpart, "Stealing machine learning models via prediction APIs," in *Proc. USENIX Security Symp.*, 2016.
9. N. Carlini et al., "Extracting training data from large language models," in *Proc. USENIX Security Symp.*, 2021. Link
10. M. Nasr et al., "Scalable extraction of training data from (production) language models," arXiv:2311.17035; presented at ICLR 2025. Link
11. K. Greshake et al., "Not what you've signed up for: Compromising real-world LLM-integrated applications with indirect prompt injection," in *Proc. ACM Workshop on Artificial Intelligence and Security (AISec)*, 2023.
12. OWASP GenAI Security Project, *OWASP Top 10 for LLM Applications 2025 (v2.0)*, OWASP Foundation, Nov. 2024. genai.owasp.org
13. R. Sommer and V. Paxson, "Outside the closed world: On using machine learning for network intrusion detection," in *Proc. IEEE Symp. Security and Privacy*, 2010.
14. M. Brundage et al., "The malicious use of artificial intelligence: Forecasting, prevention, and mitigation," arXiv:1802.07228, 2018.

Predictive AI for Disaster Risk Forecasting and Early Warning Systems

Harshit Kriti Raj Mandal[1], Gourav Gupta[2], Mrs. Apurba Sinha[3]
[1]Computer Science, G.G.S.E.S.T.C, Kandra, Bokaro, Jharkhand India; harshitkirtiraj@gmail.com
[2]Computer Science, G.G.S.E.S.T.C, Kandra, Bokaro, Jharkhand India; imguptagourav@gmail.com
[3]Assistant. Professor, Computer Science, G.G.S.E.S.T.C, Kandra, Bokaro, Jharkhand India; m.sinha0686@gmail.com

Abstract

Disaster losses keep rising because hazards, exposure and vulnerability are all changing at once, while warning lead times and local relevance have improved unevenly across regions. This paper examines how predictive artificial intelligence (AI) can strengthen the forecasting and warning chain, from observation and hazard prediction to impact estimation, communication and community response. We review the main model families (recurrent, convolutional, graph-based, transformer, generative and physics-informed networks), assess their evidence base for floods, tropical and extratropical storms, precipitation now casting, earthquakes, wildfires and landslides, and identify where each approach adds value over conventional physics-based and statistical systems. We then propose a five-layer reference framework that couple's hybrid AI-physics forecasting with probabilistic impact modeling, calibrated alerting and a feedback loop for continuous verification. The paper closes with an evaluation protocol built on proper scoring rules and event-based skill measures, and a discussion of data scarcity, explain ability, equity and governance issues that decide whether AI-enabled warnings actually save lives. This is a review and conceptual-framework study; it reports no new experimental results.

Keywords: disaster risk reduction; early warning systems; machine learning; deep learning; probabilistic forecasting; flood forecasting; now casting; explainable AI; Sendai Framework

1. Introduction

Early warning is among the most cost-effective tools for reducing disaster losses. Analyses compiled for the Global Assessment Report indicate that even a day of advance notice can cut damage from an impending hazard by a substantial margin, and that investment in hydro-meteorological services pays back several times over (Rogers & Tsirkunov, 2010). The Sendai Framework for Disaster Risk Reduction 2015-2030 therefore sets a global target to increase the availability of and access to multi-hazard early warning systems and risk information (UNDRR, 2015), and the United Nations' Early Warnings for All initiative aims to extend such coverage to everyone by the end of 2027 (United Nations, 2022).

Forecasting quality is the technical core of any warning system. For decades, that role has been played by numerical models that integrate the governing equations of the atmosphere, ocean, land surface or solid earth, supported by statistical post-processing (Bauer et al., 2015). These systems are powerful but expensive to run, slow to improve, and often poorly resolved at the scale where decisions are made: a river reach, a hillside, a neighborhood. Meanwhile, the volume of Earth observation data, from satellites, radar, seismic networks, gauges and crowd-sourced reports, has outgrown what traditional workflows can exploit (Reichstein et al., 2019).

Machine learning has begun to close this gap. Data-driven global weather models now match or exceed leading operational systems on many verification measures at a small fraction of the computational cost (Lam et al., 2023; Bi et al., 2023; Price et al., 2025), deep generative models improve short-range rainfall prediction (Ravuri et al., 2021; Zhang et al., 2023), and neural hydrological models have produced flood forecasts for river basins that have no gauges at all (Nearing et al., 2024). Yet impressive benchmark skill does not by itself make a warning system. Warnings have to be timely, calibrated, understandable, trusted and acted upon (Basher, 2006), and the literature on AI for disasters is often organized by algorithm rather than by the needs of that full chain.

This paper takes the chain as its organizing principle. Its contributions are: (i) a hazard-by-hazard synthesis of what predictive AI can and cannot yet do; (ii) a five-layer reference framework that places AI components inside the people-centered early warning model; (iii) an evaluation protocol that links forecast verification to decision value; and (iv) a candid appraisal of the technical, ethical and institutional obstacles to deployment.

2. Foundations of Early Warning Systems

An early warning system is more than a forecast. The widely used people-centered model describes four interacting elements: risk knowledge, monitoring and warning service, dissemination and communication, and response capability (Basher, 2006;

Kelman & Glantz, 2014). A failure in any element, whether a missed detection, an unintelligible message or an evacuation route that floods first, defeats the whole system.

Predictive AI can contribute to each element, but its role differs:

1. **Risk knowledge:** learning susceptibility and exposure patterns from imagery, terrain and census data, for example landslide susceptibility maps (Reichenbach et al., 2018).
2. **Monitoring and warning service:** forecasting hazard intensity, timing and location, and quantifying uncertainty.
3. **Dissemination:** selecting audiences, channels and languages, and translating hazard forecasts into expected impacts.
4. **Response capability:** supporting decisions such as reservoir release, evacuation sequencing and resource pre-positioning.

A useful distinction runs between *hazard forecasting* (how strong, where, when) and *impact-based forecasting* (what the hazard will do to people, assets and services). Modern practice is moving toward the second, because decision makers act on consequences, not on millimeters of rain. Predictive AI is especially attractive here because impact models must combine heterogeneous inputs that physics alone does not link, such as building stock, road networks, population movement and past damage records.

Two properties of the problem shape every modeling choice. First, the events that matter most are rare, so training data for extremes are scarce and verification samples are small. Second, the cost of errors is asymmetric: a missed event and a false alarm carry very different consequences, and repeated false alarms erode public trust over time.

3. Predictive AI Methods

No single architecture suits every hazard. Table 1 summarizes the main families, the data they suit and the main caveat for warning applications.

Model family	Typical hazard use	Strength	Main limitation
Tree ensembles (random forest, gradient boosting)	Susceptibility mapping, damage and impact classification	Robust on tabular data, fast, easy to interpret with SHAP (Lundberg & Lee, 2017)	Weak at spatio-temporal dynamics
Recurrent networks (LSTM, GRU)	River discharge, drought and reservoir forecasting	Captures memory and lag, generalizes across basins (Kratzert et al., 2018, 2019)	Limited spatial context; data-hungry
Convolution networks (CNN, U-Net, ConvLSTM)	Radar now casting, flood and burn-scar mapping	Learn spatial patterns from imagery (Ronneberger et al., 2015; Shi et al., 2015)	Blurring of extremes at longer lead times
Graph neural networks	Global weather, river networks, infrastructure cascades	Respect irregular geometry and connectivity (Lam et al., 2023)	Costly training; complex design
Transformers	Medium-range weather, seismic phase picking	Long-range dependencies, scalable (Bi et al., 2023; Mousavi et al., 2020)	Large compute and data needs
Generative models (diffusion, GAN)	Ensemble weather, precipitation now casting	Sharp, physically plausible fields and calibrated ensembles (Ravuri et al., 2021; Price et al., 2025)	Hard to verify; risk of plausible but wrong detail
Physics-informed and hybrid models	Hydrodynamics, climate and weather with learned components	Embed conservation laws and constraints (Raissi et al., 2019; Kochkov et al., 2024)	Training difficulty; stiff dynamics

Table 1. Model families for disaster forecasting.

3.1 Uncertainty quantification

A warning without an uncertainty estimate invites over- or under-reaction. Probabilistic output can be obtained through ensembles of independently trained networks (Lakshminarayanan et al., 2017), approximate Bayesian inference (Gal & Ghahramani, 2016) or generative sampling. Whatever the route, the resulting probabilities should be checked for reliability and recalibrated where needed, since an overconfident model is more dangerous than a modestly less accurate but honest one.

3.2 Hybrid AI-physics modeling

Pure data-driven models can break physical constraints and degrade out of distribution, while pure physics models carry structural error and high cost. Hybrid designs combine the two: learned parameterizations inside a physical solver, AI emulators that speed up expensive simulations, or machine-learning post-processing that corrects systematic bias in physical forecasts (Reichstein et al., 2019; Kochkov et al., 2024). For warning systems that must be trusted in unprecedented conditions, hybrids currently offer the best balance of skill and credibility.

3.3 Data sources

The enabling data layer includes reanalysis products such as ERA5 (Hersbach et al., 2020), large-sample hydrological archives such as CAMELS (Addor et al., 2017), labeled flood imagery (Bonafilia et al., 2020), and cloud platforms for processing satellite archives (Gorelick et al., 2017). Benchmarks such as Weather Bench make model comparison more reproducible (Rasp et al., 2020).

4. Hazard-Specific Applications

4.1 Floods and hydrological extremes

Hydrology offers the clearest evidence that AI can improve warnings. LSTM-based rainfall-runoff models outperform several established conceptual models on large benchmark sets (Kratzert et al., 2018), and models trained jointly across many catchments transfer knowledge to basins with little or no local data (Kratzert et al., 2019). Operational-scale work has since shown skilful forecasts of extreme river events in ungauged locations, with reliability at multi-day lead times comparable to that of conventional now casts for gauged ones (Nearing et al., 2024). Earlier reviews of machine learning in flood prediction already noted the rapid shift from shallow models to deep architectures (Mosavi et al., 2018; Sit et al., 2020). Satellite radar combined with segmentation networks adds event-time flood extent mapping (Bonafilia et al., 2020).

4.2 Storms and precipitation now casting

Graph and transformer weather models produce global ten-day forecasts in seconds on a single accelerator, with cyclone track skill that rivals leading numerical systems (Lam et al., 2023; Bi et al., 2023). Earlier neural approaches such as FourCastNet showed that high-resolution global emulation was feasible (Pathak et al., 2022), and probabilistic generative models now deliver ensemble forecasts that are better calibrated than their single-trajectory predecessors (Price et al., 2025). For the zero-to-two-hour window, generative radar now casting has been rated by forecasters as more useful than physics-based and earlier deep-learning baselines in a majority of cases (Ravuri et al., 2021), and later physics-conditioned models target extreme rainfall specifically (Zhang et al., 2023). Climate-scale signals also benefit: deep learning has extended skilful ENSO prediction to lead times beyond a year (Ham et al., 2019).

4.3 Earthquakes

Earthquakes cannot yet be predicted in time and place with usable reliability, so AI contributes elsewhere in the chain. Deep networks detect and pick seismic phases faster and more sensitively than classical algorithms, expanding catalogues and shortening the time to issue shaking alerts (Mousavi et al., 2020; Mousavi & Beroza, 2022). Applications of machine learning in seismology are broad, from event discrimination to ground-motion estimation (Kong et al., 2019). The honest framing is *detection and rapid characterization*, not prediction.

4.4 Wildfires

Machine learning supports fire danger rating, ignition and spread modelling, and burned-area mapping, drawing on weather, fuel, topography and satellite thermal data (Jain et al., 2020). Spread forecasts remain sensitive to wind and fuel inputs, so uncertainty and rapid data assimilation matter as much as the model class.

4.5 Landslides

Statistical and machine-learning susceptibility models are well established (Reichenbach et al., 2018). Temporal early warning, which needs rainfall thresholds and slope-monitoring data combined in near real time, is less mature, and the main limitation is the scarcity of dated landslide inventories rather than model capacity.

5. Proposed Reference Framework: AI-Enabled Multi-Hazard Early Warning

We propose a five-layer framework that maps AI components onto the people-centered warning chain. Each layer has a defined input, output and verification duty, so that weaknesses can be located rather than hidden inside one opaque model.

1. **Observation and data-assimilation layer.** Ingests gauges, radar, satellite, seismic, crowd-sourced and social-sensor streams. AI performs quality control, gap filling, super-resolution and anomaly detection. Output: a harmonized, uncertainty-tagged state estimate.
2. **Hazard prediction layer.** Runs hybrid AI-physics ensembles per hazard, using data-driven emulators for speed and physical models for constraint and out-of-distribution checks. Output: probabilistic hazard fields (depth, wind, shaking, burn probability) at several lead times.
3. **Impact and risk layer.** Converts hazard fields into expected consequences by combining exposure and vulnerability data with learned damage functions. Output: impact-based forecasts such as people affected, roads cut and facilities at risk, each with credible intervals.
4. **Decision and communication layer.** Applies cost-loss decision rules and thresholds tuned to local tolerance for false alarms, then drafts tailored multilingual messages for each audience and channel, with human sign-off for high-consequence alerts.
5. **Verification and learning layer.** Logs every forecast, warning, action and outcome, computes skill and reliability, detects drift, and triggers retraining. Community feedback and post-event surveys enter here.

Three design principles cut across the layers. *Human-in-the-loop by default:* forecasters retain authority to override automated warnings, because accountability for public alerts cannot be delegated to a model. *Graceful degradation:* when connectivity, sensors or the AI service fail, the system falls back to simpler physical or threshold rules rather than going silent. *Local adaptation:* models are fine-tuned and thresholds co-designed with the communities who will receive the alerts, since a technically skilful warning that people neither understand nor trust will not change behavior (Basher, 2006).

6. Evaluation, Challenges and Ethics

6.1 Evaluation protocol

Average error alone hides what matters for warnings. We recommend reporting three groups of measures together:

1. **Probabilistic quality:** strictly proper scoring rules such as the continuous ranked probability score and the Brier score (Gneiting & Raftery, 2007), plus reliability diagrams.
2. **Event-based skill:** probability of detection, false-alarm ratio and critical success index at the operational warning threshold, reported by lead time and for the rarest events separately.
3. **Decision value:** expected expense under a cost-loss model, and, where data exist, observed outcomes such as lead time achieved, evacuation compliance and casualties avoided.

Evaluation should use temporally and spatially blocked splits, test on held-out extreme events and ungauged regions, and compare against strong operational baselines, not weak climatology.

6.2 Technical challenges

4. **Rare extremes and non-stationary.** Models trained on historical records may fail in conditions outside past experience, which climate change makes more frequent.
5. **Data scarcity and bias.** Observation networks are sparse in many of the most exposed regions, so models can perform best where least needed. Transfer learning and global training help but do not remove the problem (Nearing et al., 2024).
6. **Explainability.** Attribution methods help forecasters sanity-check predictions (Lundberg & Lee, 2017), but explanations of black-box models can mislead, and inherently interpretable designs are preferable for high-stakes choices (Rudin, 2019).
7. **Physical consistency.** Purely statistical emulators can violate conservation laws, motivating hybrid designs (Reichstein et al., 2019).
8. **Operational robustness.** Latency, sensor outages, adversarial or corrupted inputs and model drift must be monitored continuously.

6.3 Ethical and governance issues

Warnings are public-safety infrastructure, so equity and accountability are first-order requirements. Coverage gaps by income, language, disability or connectivity can mean AI widens, rather than narrows, protection gaps. Social-media and mobile-phone data raise privacy concerns and should be used in aggregated, consented or anonymized form. Responsibility for automated alerts must stay with named institutions, with transparent documentation of model scope and known failure modes, and with independent audit. Finally, the energy and data costs of large models should be weighed against the lighter-weight alternatives that low-resource agencies can actually run (Rolnick et al., 2022).

7. Future Directions and Conclusion

7.1 Future directions

9. **Foundation models for Earth observation and hazards,** fine-tuned for local tasks with few labels.
10. **Compound and cascading hazards,** such as heavy rain on burned slopes or earthquake-triggered landslides, which single-hazard models treat separately.
11. **Impact-based and anticipatory action,** linking forecasts directly to pre-agreed funding and response triggers.
12. **Data assimilation with learned models,** so AI forecasts are continually corrected by fresh observations.
13. **Edge and low-bandwidth deployment,** enabling lightweight models and offline alerts where connectivity is poor.
14. **Open benchmarks for extremes,** including standard held-out catastrophic events and impact-oriented metrics, to prevent progress being judged on average days.

7.2 Conclusion

Predictive AI has moved from promising research to a credible component of operational forecasting, with the strongest evidence in hydrology, global weather and precipitation now casting, and a supporting rather than predictive role for earthquakes. Its value for disaster risk reduction, however, depends less on model accuracy than on how well it is embedded in the full warning chain: calibrated uncertainty, impact-based translation, human oversight, equitable reach and continuous verification. The five-layer framework proposed here offers a structure for that embedding, and the evaluation protocol links skill to decisions. Future work should test the framework in pilot deployments, report results on extreme and ungauged cases, and develop governance practices that keep AI-enabled warnings accountable to the communities they serve.

References

1. Addor, N., Newman, A. J., Mizukami, N., & Clark, M. P. (2017). The CAMELS data set: catchment attributes and meteorology for large-sample studies. *Hydrology and Earth System Sciences*, 21, 5293-5313.
2. Basher, R. (2006). Global early warning systems for natural hazards: systematic and people-centered. *Philosophical Transactions of the Royal Society A*, 364, 2167-2182.
3. Bauer, P., Thorpe, A., & Brunet, G. (2015). The quiet revolution of numerical weather prediction. *Nature*, 525, 47-55.
4. Bi, K., Xie, L., Zhang, H., Chen, X., Gu, X., & Tian, Q. (2023). Accurate medium-range global weather forecasting with 3D neural networks. *Nature*, 619, 533-538.
5. Bonafilia, D., Tellman, B., Anderson, T., & Issenberg, E. (2020). Sen1Floods11: a georeferenced dataset to train and test deep learning flood algorithms for Sentinel-1. *CVPR Workshops*.
6. Gal, Y., & Ghahramani, Z. (2016). Dropout as a Bayesian approximation: representing model uncertainty in deep learning. *ICML*.
7. Gneiting, T., & Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102, 359-378.
8. Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202, 18-27.
9. Ham, Y.-G., Kim, J.-H., & Luo, J.-J. (2019). Deep learning for multi-year ENSO forecasts. *Nature*, 573, 568-572.

10. Hersbach, H., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146, 1999-2049.
11. Jain, P., Coogan, S. C. P., Subramanian, S. G., Crowley, M., Taylor, S., & Flannigan, M. D. (2020). A review of machine learning applications in wildfire science and management. *Environmental Reviews*, 28, 478-505.
12. Kelman, I., & Glantz, M. H. (2014). Early warning systems defined. In *Reducing Disaster: Early Warning Systems for Climate Change*. Springer.
13. Kochkov, D., et al. (2024). Neural general circulation models for weather and climate. *Nature*, 632, 1060-1066.
14. Kong, Q., Trugman, D. T., Ross, Z. E., Bianco, M. J., Meade, B. J., & Gerstoft, P. (2019). Machine learning in seismology: turning data into insights. *Seismological Research Letters*, 90, 3-14.
15. Kratzert, F., Klotz, D., Brenner, C., Schulz, K., & Herrnegger, M. (2018). Rainfall-runoff modelling using long short-term memory (LSTM) networks. *Hydrology and Earth System Sciences*, 22, 6005-6022.
16. Kratzert, F., et al. (2019). Towards learning universal, regional, and local hydrological behaviors via machine learning applied to large-sample datasets. *Hydrology and Earth System Sciences*, 23, 5089-5110.
17. Lakshminarayanan, B., Pritzel, A., & Blundell, C. (2017). Simple and scalable predictive uncertainty estimation using deep ensembles. *NeurIPS*.
18. Lam, R., et al. (2023). Learning skillful medium-range global weather forecasting. *Science*, 382, 1416-1421.
19. Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *NeurIPS*.
20. Mosavi, A., Ozturk, P., & Chau, K. (2018). Flood prediction using machine learning models: literature review. *Water*, 10(11), 1536.
21. Mousavi, S. M., Ellsworth, W. L., Zhu, W., Chuang, L. Y., & Beroza, G. C. (2020). Earthquake transformer: an attentive deep-learning model for simultaneous earthquake detection and phase picking. *Nature Communications*, 11, 3952.
22. Mousavi, S. M., & Beroza, G. C. (2022). Deep-learning seismology. *Science*, 377, eabm4470.
23. Nearing, G., et al. (2024). Global prediction of extreme floods in ungauged watersheds. *Nature*, 627, 559-563.
24. Pathak, J., et al. (2022). FourCastNet: a global data-driven high-resolution weather model using adaptive Fourier neural operators. *arXiv:2202.11214*.
25. Price, I., et al. (2025). Probabilistic weather forecasting with machine learning. *Nature*, 637, 84-90.
26. Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks. *Journal of Computational Physics*, 378, 686-707.
27. Rasp, S., Dueben, P. D., Scher, S., Weyn, J. A., Mouatadid, S., & Thuerey, N. (2020). WeatherBench: a benchmark data set for data-driven weather forecasting. *Journal of Advances in Modeling Earth Systems*, 12, e2020MS002203.
28. Ravuri, S., et al. (2021). Skilful precipitation nowcasting using deep generative models of radar. *Nature*, 597, 672-677.
29. Reichenbach, P., Rossi, M., Malamud, B. D., Mihir, M., & Guzzetti, F. (2018). A review of statistically-based landslide susceptibility models. *Earth-Science Reviews*, 180, 60-91.
30. Reichstein, M., et al. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, 566, 195-204.
31. Rogers, D., & Tsirkunov, V. (2010). Costs and benefits of early warning systems. Global Assessment Report on Disaster Risk Reduction, background paper. UNISDR/World Bank.
32. Rolnick, D., et al. (2022). Tackling climate change with machine learning. *ACM Computing Surveys*, 55(2), 1-96.
33. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: convolutional networks for biomedical image segmentation. *MICCAI*.
34. Rudin, C. (2019). Stop explaining black box machine learning models for high-stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1, 206-215.

Machine Learning-Based Detection of Phishing Websites for Protecting User Data Privacy: A Calibrated and Interpretable Random-Forest Framework

Aariz Abdullah[1], Faiza Afreen[1*], David Kumar Pandey[1], Rani Perween[1], K Shreya Rao[1]

[1]Department of Computer Science and Engineering, Al Kabir Institute of Management and Technology, Jamshedpur, Jharkhand, India

*Corresponding author e-mail: btcse007.25@gmail.com

Abstract

Background: Phishing websites imitate genuine services to obtain passwords, banking details and identity data, and about 3.8 million phishing attacks were recorded in 2025. Blacklists cannot recognise newly created pages, while recent learning-based detectors are often complex ensembles evaluated on a single split of a benchmark that contains many duplicate records. **Objectives:** The study aims to develop a simple and interpretable machine learning detector that identifies phishing websites from observable URL, domain and page characteristics, and to evaluate it rigorously against recent methods. **Methodology:** The UCI Phishing Websites dataset (11,055 records, 30 features) was used with stratified 80:20 splits repeated ten times. Logistic regression, decision tree, support vector machine and random forest were compared with six detectors published between 2020 and 2026, all tuned with the same budget, using accuracy, precision, recall, F1-score, Wilcoxon and Friedman tests, a duplicate-aware split and a second dataset of 10,000 pages. **Proposed model:** The framework combines a data audit, a tuned random forest, a validation-calibrated decision threshold and TreeSHAP explanations, and is integrated into a web interface that returns a safe or phishing verdict with its reasons. **Results:** Random forest was the most accurate classical classifier (97.18 ± 0.28%), ahead of the support vector machine (96.78%), decision tree (96.41%) and logistic regression (92.72%). The calibrated detector reached 97.07% accuracy, 96.95% precision, 96.69% F1-score and the highest phishing recall of all eleven methods (96.47%); the best stacking ensemble was 0.33 points more accurate but required about eight times longer training. SSL state, anchor URL and web traffic were the most influential indicators. A duplicate-aware split reduced accuracy to 95.26%, and the detector reached 98.38% on the second dataset. **Conclusion:** A single calibrated forest attains the performance band of recent ensembles with transparent reasoning, and conventional evaluation on the UCI benchmark overstates accuracy by about 1.8 points. **Real-time applications:** Scoring takes about 10 ms per URL on one CPU thread, so the detector suits browser extensions, e-mail gateways and campus proxies that protect personal data of everyday users.

Keywords: cyber security; data privacy; phishing detection; machine learning; random forest; explainable AI

1. Introduction

Phishing websites imitate trusted services to obtain passwords, banking credentials and identity documents. The Anti-Phishing Working Group recorded about 3.8 million phishing attacks in 2025, slightly more than in 2024 [1], and each successful attack is a direct and often irreversible loss of personal data.

Browser blacklists flag a page only after it has been reported and verified, whereas phishing pages are short-lived and cheaply regenerated, so new pages are often visited before they are listed [2], [3]. Learning-based detectors classify a page from observable URL, domain and HTML properties, but three weaknesses persist. First, many recent detectors are stacking or voting ensembles whose gains over a single tuned forest are small and rarely accompanied by repeated-run variance, significance tests or inference cost [4]–[6]. Second, the widely used UCI Phishing Websites benchmark contains many exact duplicates, so random splits place identical vectors in training and test data, and the resulting optimism is seldom quantified. Third, most studies fix the decision threshold at 0.5 and use a single dataset.

The contributions of this study are as follows.

5. A phishing detection framework is proposed that combines a data audit, an equal-budget tuned random forest, a validation-calibrated decision threshold and TreeSHAP explanations. On the UCI dataset (11,055 records, 30 features) it reaches 97.07 ± 0.40% accuracy and the highest phishing recall of eleven methods (96.47%), 4.35 percentage points above logistic regression and within 0.33 points of the most accurate stacking ensemble, while training about eight times faster.
6. Six detectors published between 2020 and 2026 are re-implemented and tuned with an identical budget, and all comparisons are reported as mean ± standard deviation over repeated runs with Wilcoxon signed-rank and Friedman tests.
7. A duplicate-aware evaluation, in which identical feature vectors never appear on both sides of the split, shows that accuracy on genuinely unseen vectors is 95.26%, about 1.8 points below the conventional estimate, and identifies the feature groups that carry the decision.
8. Generalisation is examined on an independent dataset (Mendeley, 10,000 records, 48 features), where the detector reaches 98.38% accuracy, and its latency, memory and deployment cost are reported.

2. Related Work

List-based and heuristic detection. Blacklists give few false positives but miss zero-day pages, and manual heuristics age quickly as attackers adapt [2], [3]. Mohammad et al. formalised thirty heuristics and released them as the UCI Phishing Websites dataset, still the most used benchmark [7], [8].

Classical machine learning. Random forest, decision tree, SVM and logistic regression trained on engineered URL and page features are the most widely studied detectors [9], with accuracies on the UCI data typically between about 92% and 97% [5], [10], [16]. Alnemari and Alshammari found the forest the most accurate of four models [10], Kapan and Gunal found that URL and HTTP features performed better than the full feature set [11], and Sahingoz et al. reached high accuracy from lexical URL features without third-party services [12]. Variance across repeated splits is seldom reported.

Ensembles and optimised boosting. Zamir et al. stacked a random forest, a neural network and bagging [4]; Al-Sarem et al. tuned a stacking ensemble with a genetic algorithm [5]; Karim et al. combined LR, SVM and DT by voting [13]; Bahaghighat et al. and Sheikhi and Kostakos used tuned XGBoost, the latter with particle swarm optimisation [14], [15]; PhishGuard combined feature selection with a multi-layer stacking ensemble and a CatBoost meta-learner [6]; and EnLeM combined DT, RF and KNN by hard voting, reporting 97.21% on the UCI data [16]. Reported gains over a single forest are usually below one point, and the added cost is rarely measured.

Deep and raw-content models. Aljofey et al. combined character-level URL and HTML features in XGBoost [17], and Opara et al. applied a convolutional network to raw URL and HTML embeddings [18]. Tree ensembles nevertheless remain highly competitive on tabular data [19], and results depend strongly on benchmark construction [20].

Research gap. Table 1 compares representative detectors. None of them combines a single interpretable model, threshold calibration and a duplicate-aware leakage check, and repeated-run significance tests are not reported.

Table 1. Properties of representative phishing detectors, as described in the cited works (✓ = present, – = absent or not reported).

Study	Model	Single model at inference	No meta-learner	Native tree explanation	Calibrated threshold	Leakage check	Second dataset
Zamir et al. 2020 [4]	Stacking (RF, NN, bagging)	–	–	–	–	–	–
Al-Sarem et al. 2021 [5]	GA-optimised stacking	–	–	–	–	–	✓
Alnemari & Alshammari 2023 [10]	ANN, SVM, DT, RF	✓	✓	✓	–	–	–

Study	Model	Single model at inference	No meta-learner	Native tree explanation	Calibrated threshold	Leakage check	Second dataset
Sheikhi & Kostakos 2024 [15]	PSO-tuned XGBoost	✓	✓	✓	–	–	–
Ovi et al. 2024 [6]	Stacking, CatBoost meta	–	–	–	–	–	✓
Yeasmin et al. 2026 [16]	Hard-voting DT, RF, KNN	–	✓	–	–	–	–
This study	Tuned RF + calibrated threshold	✓	✓	✓	✓	✓	✓

3. Materials and Methods

3.1 Problem formulation

Let $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^{n}$ denote a labelled collection of websites, where $\mathbf{x}_i \in \{-1,0,1\}^d$ encodes $d = 30$ ternary indicators of the URL, domain and page content and $y_i \in \{0,1\}$ equals 1 for a phishing page. The task is to learn a scoring function $p: \{-1,0,1\}^d \to [0,1]$ and a threshold τ such that the decision rule

$$\hat{y}(\mathbf{x}) = \mathbb{1}[p(\mathbf{x}) \geq \tau] \qquad (1)$$

minimises the expected misclassification cost on unseen websites. Because the true distribution is unknown, the hyperparameters $\boldsymbol{\theta}$ of the scoring model are selected by K-fold cross-validation on the training split $\mathcal{D}_{tr}$,

$$\boldsymbol{\theta}^* = \underset{\boldsymbol{\theta} \in \Theta}{\operatorname{argmax}} \frac{1}{K} \sum_{k=1}^{K} \operatorname{Acc}\left(p_{\boldsymbol{\theta}}; \mathcal{D}_{tr}^{(k)}\right), \qquad (2)$$

and the threshold is selected on a held-out validation subset $\mathcal{D}_{val} \subset \mathcal{D}_{tr}$ that is never used for fitting the final forest. Three assumptions are made: (i) training and deployment websites are drawn from the same feature distribution; (ii) the 30 indicators can be computed for a URL at query time; and (iii) misclassifying a phishing page and a legitimate page carry comparable cost, so the F1-score is an appropriate selection criterion. The principal notation is: n, number of records; d, number of features; T, number of trees; m, features considered per split; h_t, the t-th tree; τ, decision threshold; ϕ_j, SHAP value of feature j.

3.2 Proposed framework

The framework (Figure 1) has five modules in an offline training path and an online inference path, each responding to a weakness identified in Section 1.

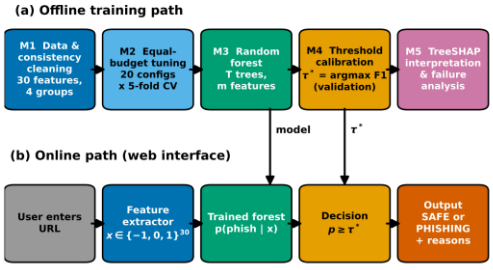


Figure 1. Overview of the proposed phishing detection framework. (a) Offline path: data audit (M1), equal-budget hyperparameter tuning (M2), random forest training (M3), validation-based threshold calibration (M4) and TreeSHAP interpretation (M5). (b) Online path used by the web interface: the URL is converted to the 30-feature vector, scored by the trained forest and compared with the calibrated threshold τ^*; the most influential features are returned with the verdict.

M1: Data audit. The duplicate and label-conflict structure of the data is measured before training. A record i is a conflict if another record shares its feature vector but carries the opposite label,

$$\mathcal{C} = \{\, i : \exists j \neq i,\ \mathbf{x}_j = \mathbf{x}_i,\ y_j \neq y_i \,\}. \qquad (3)$$

It motivates the duplicate-aware evaluation of Section 4.5 but does not alter the training data in the main experiments, keeping results comparable with published baselines.

M2: Equal-budget tuning. Every learner, including the baselines, receives 20 random configurations evaluated by five-fold cross-validation (Eq. 2), so that gains of complex ensembles cannot stem from unequal tuning effort.

M3: Random forest. The forest suits discrete, interacting indicators and provides probabilities and exact attributions at low cost. Each tree is grown on a bootstrap sample and, at every node S, selects among m randomly drawn features the split that maximises the decrease in Gini impurity,

$$G(S) = 1 - \sum_{c\in\{0,1\}} \pi_c\,(S)^2, \qquad (4)$$

$$\Delta G = G(S) - \frac{|S_L|}{|S|}G(S_L) - \frac{|S_R|}{|S|}G(S_R), \qquad (5)$$

where $\pi_c(S)$ is the proportion of class c in node S. The phishing probability is the average of the T tree votes,

$$p(\mathbf{x}) = \frac{1}{T}\sum_{t=1}^{T} h_t\,(\mathbf{x}), \qquad h_t(\mathbf{x}) \in [0,1]. \qquad (6)$$

Proposition 1 (generalisation of the forest). Let s denote the strength of the forest, that is, the expected margin of the correct class, and $\bar{\rho}$ the mean correlation between the margin functions of two trees. Breiman [21] showed that the generalisation error satisfies

$$PE^* \le \frac{\bar{\rho}\,(1-s^2)}{s^2}. \qquad (7)$$

The bound decreases as trees become less correlated, which is the role of bootstrap sampling and of restricting each split to $m < d$ features. It also implies that, once T is large enough for the average in Eq. (6) to stabilise, adding trees no longer changes the error; this behaviour is checked empirically in Section 4.4.

M4: Threshold calibration. A fixed threshold of 0.5 is not necessarily optimal for the F1-score [22]. The threshold is therefore chosen on the validation subset from a grid $\mathcal{T} = \{0.30, 0.31, \ldots, 0.70\}$ of $|\mathcal{T}| = 41$ values,

$$\tau^* = \operatorname*{argmax}_{\tau\in\mathcal{T}} \mathrm{F1}(\tau; \mathcal{D}_{val}). \qquad (8)$$

Proposition 2 (reliability of the selected threshold). For any fixed τ, the validation accuracy is an average of n_v independent bounded losses. By Hoeffding's inequality and a union bound over the finite grid, with probability at least $1-\delta$ the validation accuracy of every candidate threshold, including τ^*, lies within

$$\varepsilon = \sqrt{\frac{\ln(2|\mathcal{T}|/\delta)}{2n_v}} \qquad (9)$$

of its expected value. With $n_v = 885$ validation records, $|\mathcal{T}| = 41$ and $\delta = 0.05$, $\varepsilon \approx 0.065$, and the effect of calibration is therefore quantified empirically in the ablation study (Section 4.2).

M5: Interpretation. For a prediction $p(\mathbf{x})$, the SHAP value of feature j distributes the prediction among features according to the Shapley value,

$$\phi_j = \sum_{S\subseteq F\setminus\{j\}} \frac{|S|!\;(d-|S|-1)!}{d!}\big(f_x(S\cup\{j\}) - f_x(S)\big), \qquad (10)$$

and the attributions are additive, $p(\mathbf{x}) = \phi_0 + \sum_{j=1}^{d}\phi_j$ (11). TreeSHAP computes Eq. (10) exactly for tree ensembles in polynomial time [23], so the web interface can return the three features with the largest ϕ_j as the reasons behind each verdict.

The training and inference procedures are given in Algorithms 1 and 2. In Algorithm 1, Lines 1–2 perform the audit of M1, Lines 3–6 implement the search of Eq. (2), Lines 7–8 fit the forest on the reduced training set, Lines 9–13 select τ^* by Eq. (8), and Lines 14–15 refit the forest on all training data.

Algorithm 1. Training of the proposed detector

Input: training set $\mathcal{D}_{tr}$; search space Θ; budget $B = 20$; folds $K = 5$; grid $\mathcal{T}$
Output: forest p^*, threshold τ^*
1: hash every $\mathbf{x}_i$; $\mathcal{C} \leftarrow$ conflicting records (Eq. 3)
2: report $|\mathcal{C}|$ and the number of duplicates
3: draw $\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_B$ from Θ
4: **for** $b = 1$ to B **do** $a_b \leftarrow K$-fold accuracy of the forest with $\boldsymbol{\theta}_b$
5: **end for**
6: $\boldsymbol{\theta}^* \leftarrow \boldsymbol{\theta}_{\mathrm{argmax}_b a_b}$ (Eq. 2)
7: split $\mathcal{D}_{tr}$ into $\mathcal{D}_a$ (90%) and $\mathcal{D}_{val}$ (10%), stratified
8: grow $h_1, \dots, h_T$ on bootstraps of $\mathcal{D}_a$ (Eqs. 4–5); form p_a (Eq. 6)
9: **for each** $\tau \in \mathcal{T}$ **do**
10: $\quad \hat{y}_i \leftarrow \mathbb{1}[p_a(\mathbf{x}_i) \geq \tau]$ for $i \in \mathcal{D}_{val}$
11: $\quad F_\tau \leftarrow \mathrm{F1}(\hat{y}, y)$ on $\mathcal{D}_{val}$
12: **end for**
13: $\tau^* \leftarrow \mathrm{argmax}_\tau F_\tau$ (Eq. 8)
14: refit p^* with $\boldsymbol{\theta}^*$ on all of $\mathcal{D}_{tr}$
15: **return** p^*, τ^*

Complexity of Algorithm 1. Growing one tree on n records costs $O(m\, n\log n)$, so the forest costs $O(T\, m\, n\log n)$ and the search of Lines 4–9 costs $O(B\, K\, T\, m\, n\log n)$. Threshold selection adds $O(|\mathcal{T}|\, n_v)$. Memory is $O(T\, N_{node})$ for the forest, where N_{node} is the average number of nodes per tree, plus $O(n\, d)$ for the data.

In Algorithm 2, Lines 1–4 compute the indicators from the URL and the fetched page, Lines 5–8 traverse the trees and average their votes (Eq. 6), Lines 9–10 apply the calibrated rule of Eq. (1), and Lines 11–15 compute TreeSHAP attributions (Eq. 10) and return the three strongest reasons.

Algorithm 2. Online prediction with explanation

Input: URL u; forest $p^* = \{h_t\}_{t=1}^{T}$; threshold τ^*
Output: verdict $\hat{y}$, score p, reasons R
1: parse u into host, path and query; resolve domain
2: compute address-bar indicators $x_1, \dots, x_{12}$
3: fetch page and compute abnormal and HTML/JavaScript indicators $x_{13}, \dots, x_{23}$
4: query domain age, DNS, traffic and index indicators $x_{24}, \dots, x_{30}$
5: $s \leftarrow 0$
6: **for** $t = 1$ to T **do**
7: $\quad s \leftarrow s + h_t(\mathbf{x})$ (root-to-leaf traversal)
8: **end for**; $p \leftarrow s/T$ (Eq. 6)
9: **if** $p \geq \tau^*$ **then** $\hat{y} \leftarrow$ PHISHING
10: **else** $\hat{y} \leftarrow$ SAFE
11: compute $\phi_1, \dots, \phi_d$ by TreeSHAP (Eq. 10)
12: sort features by $|\phi_j|$ in descending order
13: $R \leftarrow$ top three features with their values
14: display $\hat{y}$, p and R to the user
15: **return** $\hat{y}, p, R$

Complexity of Algorithm 2. Prediction costs $O(T\, D)$, where D is the average tree depth, and requires no memory beyond the stored forest. TreeSHAP costs $O(T\, L\, D^2)$ for L leaves per tree; it is computed only when the user requests reasons.

3.3 Datasets, preprocessing and experimental protocol

Primary dataset. The UCI Phishing Websites dataset [8] contains 11,055 records (4,898 phishing, 6,157 legitimate) with 30 indicators coded −1, 0 or +1 [7] in four groups: address bar (G1, 12 features, e.g. IP address, URL length, sub-domains, SSL state), abnormal content (G2, 6, e.g. request URL, URL of anchor), HTML/JavaScript (G3, 5, e.g. redirection, pop-ups) and domain/third-party (G4, 7, e.g. domain age, DNS record, web traffic, page rank). There are no missing values. The audit (M1) found 5,206 exact duplicates, leaving 5,785 distinct vectors, and 357 records whose vector appears with both labels.

Second dataset. The Mendeley Phishing Dataset for Machine Learning [24] contains 10,000 balanced pages collected independently and described by 48 numeric URL and HTML features; every model is re-tuned and retrained on it.

Protocol. Each dataset is split 80:20 by stratified sampling, repeated with ten seeds on the UCI data and five on the second dataset; metrics are mean ± standard deviation. Hyperparameters are tuned once on the seed-0 training split (Table 2) and then fixed. With phishing as the positive class,

$$\mathrm{Acc} = \frac{TP + TN}{TP + TN + FP + FN}, \quad \mathrm{Pre} = \frac{TP}{TP + FP}, \qquad (12)$$

$$\mathrm{Rec} = \frac{TP}{TP + FN}, \quad \mathrm{F1} = \frac{2\,\mathrm{Pre} \cdot \mathrm{Rec}}{\mathrm{Pre} + \mathrm{Rec}}. \qquad (13)$$

together with the AUC. Per-run differences δ_r are tested with the two-sided Wilcoxon statistic $W^+ = \sum_{r:\,\delta_r > 0} \mathrm{rank}\,(|\delta_r|)$ (14), and all models jointly with the Friedman test [25],

$$\chi_F^2 = \frac{12N}{k(k+1)}\left[\sum_{j=1}^{k}\bar{R}_j^2 - \frac{k(k+1)^2}{4}\right], \qquad (15)$$

where N is the number of runs, k the number of models and $\bar{R}_j$ the mean rank of model j.

Baselines. Six detectors from 2020–2026 were re-implemented: the stacking ensemble of random forest, neural network and bagging with a logistic meta-learner (Stack-2020) [4]; a stacking of random forest, AdaBoost, XGBoost and LightGBM, a subset of the GA-optimised learners of Al-Sarem et al. (OptStack-2021) [5]; the neural network of Alnemari and Alshammari (ANN-2023) [10]; XGBoost tuned by particle swarm optimisation (PSO-XGB-2024) [15]; a PhishGuard-style stacking of gradient boosting, random forest and XGBoost under a CatBoost meta-learner (PhishGuard-2024) [6]; and the hard-voting ensemble of decision tree, random forest and k-nearest neighbours (EnLeM-2026) [16]. LR, DT, SVM and the tuned forest with a fixed 0.5 threshold (RF) complete the comparison; ensemble members reuse the tuned base settings.

Implementation. Python 3.13, scikit-learn 1.9.1, XGBoost 3.4.1, LightGBM 4.7.0, CatBoost 1.2.10 and SHAP 0.54.1 on a two-core 2.10 GHz Intel Xeon CPU with 7 GB RAM; latency is the median of 300 single-thread predictions.

Table 2. Search spaces and selected hyperparameters (UCI data, 20 random configurations, five-fold cross-validation).

Model	Search space	Selected setting
LR	C ∈ [0.001, 100], log scale	C = 0.083
DT	depth 8–30 or none; min. leaf 1–7; Gini/entropy	none; 1; entropy
SVM (RBF)	C ∈ [0.1, 100], γ ∈ [0.001, 1], log scale	C = 67.8, γ = 0.111
ANN	layers (64), (128), (128,64), (64,32); α ∈ [1e-5, 1e-2]; rate ∈ [1e-4, 1e-2]	(64,32); α = 6.2e-5; rate = 3.0e-3
PSO-XGB	rate 0.01–0.3; depth 3–10; trees 100–600 (PSO, 6 particles × 4 iterations, 3-fold CV)	0.155; 10; 408
RF / Proposed	T = 100–500; m ∈ {√d, $\log_2$d, 0.2d, 0.3d, 0.5d}; depth 20, 30, none; leaf 1–3; class weight none/balanced	T = 200, m = √d, depth 20, leaf 1
Proposed M4	τ ∈ {0.30, 0.31, …, 0.70}, 10% validation split	median τ* = 0.43

4. Results and Discussion

4.1 Comparison with recent baselines

Table 3 and Figure 2 report the UCI results. Among the classical classifiers, LR reaches 92.72%, DT 96.41%, SVM 96.78% and the tuned forest 97.18%. Stack-2020 obtains the highest accuracy (97.40%) and F1-score (97.05%), followed by PSO-XGB-2024 (97.36%) and OptStack-2021 (97.29%). The proposed detector reaches 97.07% accuracy, 96.69% F1-score and, owing to its calibrated threshold, the highest phishing recall (96.47%), i.e. the fewest phishing pages reaching a user undetected.

Table 3. Performance on the UCI dataset (mean ± standard deviation over ten runs, %). Best value in bold, second best underlined; p is the Wilcoxon signed-rank p-value of the accuracy difference with respect to the proposed detector.

Model	Accuracy	Precision	Recall	F1-score	AUC	p
LR	92.72 ± 0.38	93.17 ± 0.70	90.18 ± 0.71	91.65 ± 0.44	97.88	0.002
DT	96.41 ± 0.39	97.11 ± 0.49	94.71 ± 0.74	95.90 ± 0.45	97.45	0.004
SVM	96.78 ± 0.34	97.21 ± 0.66	95.47 ± 0.72	96.33 ± 0.39	98.66	0.008
Stack-2020 [4]	**97.40 ± 0.30**	97.79 ± 0.42	96.33 ± 0.68	**97.05 ± 0.34**	99.69	0.008
OptStack-2021 [5]	97.29 ± 0.33	97.76 ± 0.54	96.09 ± 0.63	96.92 ± 0.38	**99.71**	0.189
ANN-2023 [10]	96.50 ± 0.44	97.09 ± 0.64	94.96 ± 1.20	96.00 ± 0.53	99.45	0.002
PSO-XGB-2024 [15]	97.36 ± 0.34	97.83 ± 0.44	96.17 ± 0.63	96.99 ± 0.39	99.70	0.010
PhishGuard-2024 [6]	97.15 ± 0.44	97.21 ± 0.60	96.34 ± 0.76	96.77 ± 0.50	99.67	0.656
EnLeM-2026 [16]	97.20 ± 0.35	**98.00 ± 0.29**	95.62 ± 0.76	96.80 ± 0.41	97.04	0.215
RF (tuned, τ = 0.5)	97.18 ± 0.28	97.86 ± 0.47	95.73 ± 0.67	96.79 ± 0.32	99.68	0.207
Proposed	97.07 ± 0.40	96.95 ± 1.34	**96.47 ± 1.43**	96.69 ± 0.46	99.68	–

The Friedman test rejects equal performance ($\chi_F^2 = 78.97$, $p < 10^{-12}$). The proposed detector is significantly more accurate than LR, DT, SVM and ANN-2023 ($p \leq 0.008$), Stack-2020 and PSO-XGB-2024 are significantly more accurate than it by 0.33 and 0.29 points, and the remaining differences are not significant. The two leading methods therefore hold a small but consistent accuracy advantage on this benchmark, while the forest-based detector lies in the same performance band with the highest recall and a lower training cost (Section 4.5).

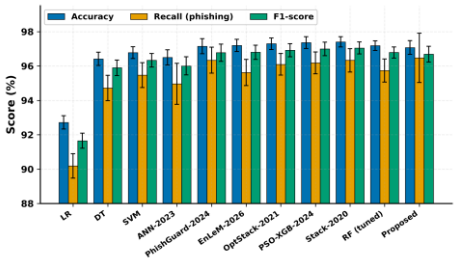


Figure 2. Accuracy, phishing recall and F1-score (mean ± standard deviation over ten runs) of all methods on the UCI dataset. Classical models are shown first and the recent detectors in increasing order of accuracy; the error bars show that the spread across runs (0.3–0.4 points) is of the same order as the differences among the six strongest methods.

In the pooled confusion matrix (Figure 3, 22,110 predictions), 346 of 9,800 phishing pages (3.53%) were missed and 301 of 12,310 legitimate pages (2.45%) were flagged. The ROC curves of the proposed detector, Stack-2020 and PSO-XGB-2024 almost coincide (AUC 99.68–99.70%, Figure 4), whereas LR, DT and SVM lie clearly below at false positive rates under 5%; the remaining gap is thus a matter of operating point rather than ranking quality.

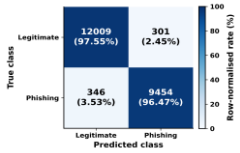


Figure 3. Confusion matrix of the proposed detector aggregated over the ten UCI test splits. Counts and row-normalised rates are shown; missed phishing pages (lower left) slightly outnumber false alarms (upper right).

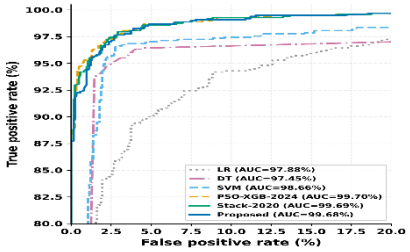


Figure 4. ROC curves (first run) of six representative methods, enlarged to false positive rates below 20%. The proposed detector, Stack-2020 and PSO-XGB-2024 have practically identical curves and AUC values.

4.2 Ablation study

Table 4 and Figure 5 remove each module and feature group in turn. Without threshold calibration (M4), accuracy rises from 97.07% to 97.17% and precision from 96.95% to 97.84%, but recall falls from 96.47% to 95.73%: the calibration chose $\tau < 0.5$ in eight of ten runs and traded about 0.9 points of precision for 0.7 points of recall, leaving F1 nearly unchanged. It therefore shifts the operating point towards fewer missed phishing pages rather than improving accuracy, consistent with Proposition 2. Default rather than tuned hyperparameters (M2) and removal of conflicting training records change accuracy by at most 0.1 point.

Feature groups matter far more: withholding G1 costs 4.79 points of accuracy and 6.81 of recall, G2 2.49 and G4 2.17 points, and G3 only 0.04 points. Since G4 requires external services, a fully local detector would lose about two points.

Table 4. Ablation results on the UCI dataset (mean ± standard deviation over ten runs, %). M2: tuning; M4: threshold calibration; G1–G4: feature groups; CR: conflict removal from the training split. Best value in bold, second best underlined.

Variant	Accuracy	Precision	Recall	F1-score
Full	97.07 ± 0.40	96.95 ± 1.34	96.47 ± 1.43	96.69 ± 0.46
− M4	**97.17 ± 0.28**	**97.84 ± 0.47**	95.73 ± 0.67	**96.78 ± 0.32**
− M2	97.15 ± 0.38	97.22 ± 1.07	96.35 ± 0.93	96.77 ± 0.42
+ CR	96.97 ± 0.28	96.60 ± 0.83	**96.57 ± 0.66**	96.58 ± 0.31
− G1	92.28 ± 0.66	92.88 ± 3.28	89.66 ± 2.95	91.15 ± 0.65
− G2	94.58 ± 0.39	93.81 ± 1.44	94.01 ± 1.04	93.90 ± 0.39
− G3	97.03 ± 0.24	97.55 ± 1.35	95.73 ± 1.56	96.62 ± 0.29
− G4	94.90 ± 0.53	95.52 ± 1.67	92.89 ± 0.94	94.17 ± 0.55

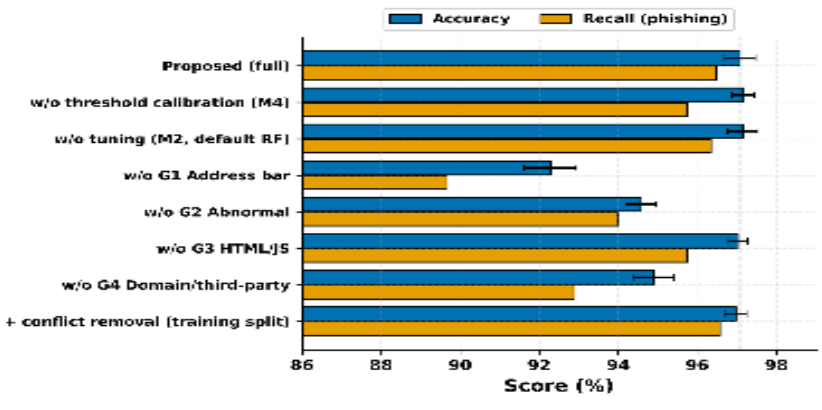


Figure 5. Ablation of modules and feature groups. Blue bars give accuracy (± standard deviation) and orange bars phishing recall; the dashed line marks the accuracy of the full detector. Removing a feature group, especially G1, has a much larger effect than removing a modelling step.

4.3 Sensitivity and convergence

In Figure 6, all 36 combinations of T and m lie between 96.82% and 97.09%, a range smaller than the run-to-run deviation, with slightly lower accuracy for $m \geq 12$ where trees become more correlated, as Eq. (7) suggests. In Figure 7 the out-of-bag error falls steeply over the first 50 trees and changes by less than 0.2 points beyond 150, so $T = 200$ lies on the stable part of the curve.

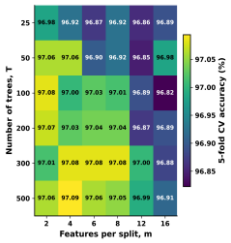


Figure 6. Sensitivity of five-fold cross-validation accuracy to the number of trees T and the number of features per split m. The surface is flat within 0.27 points, with a mild decline for $m \geq 12$.

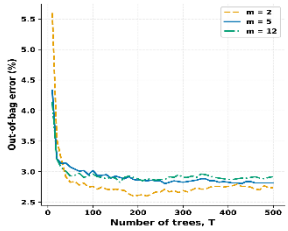


Figure 7. Out-of-bag error versus number of trees for three values of m. The error stabilises after about 150 trees, in line with Proposition 1.

4.4 Interpretability

In Figure 8 (600 test websites), the SSL state (mean $|\phi| = 0.184$) and the URL of anchor (0.155) dominate, followed by web traffic (0.059) and sub-domains (0.055), the same order as the impurity importances. A trusted certificate (+1) pushes towards legitimate, whereas a missing or untrusted certificate or mostly foreign anchors push towards phishing, in line with established heuristics [7], [12]. The interface presents these attributions as short reasons such as "no trusted HTTPS certificate".

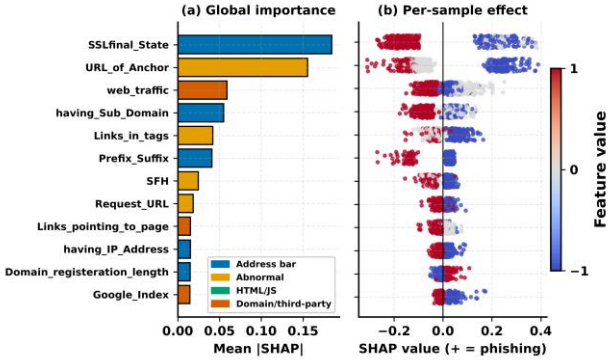


Figure 8. TreeSHAP explanation of the proposed detector. (a) Mean absolute SHAP value of the twelve most influential features, coloured by feature group. (b) SHAP value of each test website for the same features, coloured by feature value (blue = −1, grey = 0, red = +1); points to the right push the prediction towards phishing.

4.5 Generalisation, leakage check and computational cost

Table 5 summarises the generalisation and cost analyses. On the independent Mendeley dataset, all methods are more accurate than on the UCI data and the ranking is similar: PSO-XGB-2024 (98.83%) and OptStack-2021 (98.82%) lead, the tuned forest reaches 98.47% and the proposed detector 98.38 ± 0.21%, whereas LR (94.62%) and DT (96.96%) remain the weakest. With five runs the smallest attainable two-sided Wilcoxon p-value is 0.0625, so none of these pairwise differences is significant at the 5% level; the Friedman test ($\chi_F^2 = 46.36$, $p < 10^{-5}$) mainly reflects the gap between the weak classical models and the rest. With a duplicate-aware UCI split, in which each distinct vector lies entirely on one side, accuracy falls from 97.07% to 95.26 ± 0.54% and recall to 93.80%; accuracies obtained with random splits on this benchmark, as in Table 3, are thus about 1.8 points higher than those on unseen feature vectors.

Table 5. Generalisation to the Mendeley dataset (mean ± standard deviation over five runs, %) and computational cost on the UCI data (training time on 8,844 records, median single-thread latency of one URL, serialised model size). Best value in bold, second best underlined; † includes the additional forest fitted for threshold calibration.

Model	Mendeley accuracy	Mendeley F1-score	Training (s)	Latency (ms)	Size (kB)
LR	94.62 ± 0.66	94.67 ± 0.66	<u>0.03</u>	<u>1.02</u>	**3**
DT	96.96 ± 0.49	96.98 ± 0.49	**0.02**	**0.83**	<u>76</u>
SVM	97.67 ± 0.21	97.68 ± 0.20	9.18	1.14	701

Model	Mendeley accuracy	Mendeley F1-score	Training (s)	Latency (ms)	Size (kB)
Stack-2020 [4]	98.72 ± 0.17	98.72 ± 0.17	9.26	13.37	20,645
OptStack-2021 [5]	98.82 ± 0.22	98.82 ± 0.22	10.93	20.23	24,397
ANN-2023 [10]	97.32 ± 0.45	97.33 ± 0.44	0.94	1.03	106
PSO-XGB-2024 [15]	**98.83 ± 0.32**	**98.83 ± 0.32**	0.59	3.32	1,635
PhishGuard-2024 [6]	98.70 ± 0.35	98.70 ± 0.35	10.57	15.20	21,933
EnLeM-2026 [16]	98.48 ± 0.31	98.48 ± 0.31	0.61	13.50	22,223
RF (tuned, $\tau = 0.5$)	98.47 ± 0.14	98.47 ± 0.14	0.63	10.41	19,928
Proposed	98.38 ± 0.21	98.38 ± 0.22	1.17†	10.28	19,928

Figure 9 relates accuracy to single-thread latency. The forest trains in 0.63 s (1.17 s including the calibration fit), about eight times faster than Stack-2020 (9.26 s) and nine times faster than OptStack-2021 (10.93 s) even with calibration, and answers a single URL in about 10 ms, compared with 13–20 ms for the stacking ensembles. PSO-tuned XGBoost is both slightly more accurate and faster at inference (3.3 ms) with a smaller model (1.6 MB versus 19.9 MB), which makes it a competitive alternative when the particle-swarm tuning step is affordable; LR, DT and ANN answer in about 1 ms but are 0.7–4.5 points less accurate.

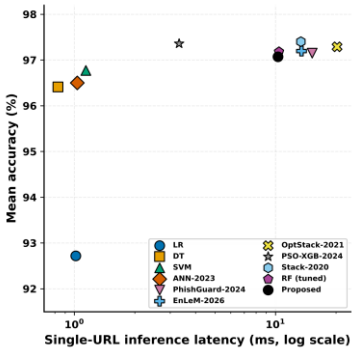


Figure 9. Mean accuracy versus single-URL inference latency (single thread, log scale) on the UCI data. Methods in the upper left offer the best accuracy per unit of computation.

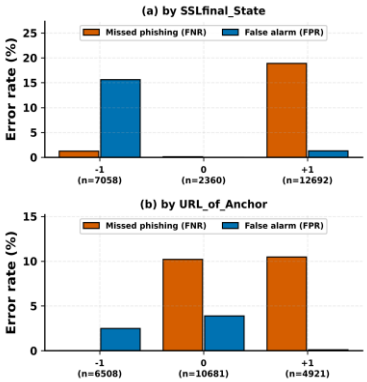


Figure 10. Error rates of the proposed detector pooled over the ten UCI test splits, broken down by the value of the two most influential features. Missed phishing pages concentrate where the SSL state is +1 and the anchor indicator is 0 or +1.

4.6 Failure-case analysis

Figure 10 shows that phishing pages with a trusted certificate (SSL state +1) are missed about 19% of the time, against about 1% otherwise, while legitimate pages without one raise false alarms about 16% of the time; phishing pages whose anchors mostly stay on their own domain are missed about 10% of the time. An attacker with a free certificate and a self-contained page removes both strongest cues, which calls for content-level evidence such as visual similarity to a target brand rather than a different classifier.

4.7 Deployment perspective

In the web interface, a request passes through feature extraction, scoring and explanation. Scoring costs about 10 ms for one URL on one CPU thread, or 0.04 ms per URL in batches, while a TreeSHAP explanation of one URL takes about 90 ms; both are small compared with fetching the page and querying domain information, which take hundreds of milliseconds over a network. The 19.9 MB forest fits easily in a browser-extension backend, a campus proxy or a low-cost cloud instance, and can be reduced by limiting tree depth if storage is constrained. No GPU is required, so the operating cost is that of an ordinary web server, and omitting the third-party indicators (G4) would allow fully local operation at a cost of about two points of accuracy (Table 4).

4.8 Limitations and threats to validity

Data age: the UCI data are more than a decade old, and several third-party indicators have since changed or been discontinued; the G4 ablation quantifies the effect of losing them. Data leakage: random splits on this benchmark give accuracies about 1.8 points higher than duplicate-aware splits, an effect that applies equally to all compared methods. Tuning and baselines: hyperparameters were tuned on the first split and then fixed for all methods, and the baselines follow their published descriptions. Statistical power: the threshold is selected on about 885 validation records, which by Proposition 2 resolves differences of several points, and five runs on the second dataset limit the attainable Wilcoxon p-value to 0.0625. Concept drift: attackers adapt to published features, so periodic retraining is required. Overfitting was controlled by cross-validated tuning on training data only, an untouched test split and repeated runs.

5. Conclusion

Under equal tuning, repeated runs and significance tests, a calibrated random-forest detector was compared with six recent detectors. On the UCI dataset the tuned forest reached 97.18% accuracy, the highest of the four classical classifiers, and the threshold-calibrated detector reached 97.07% accuracy with the highest phishing recall (96.47%). The best stacking ensemble was more accurate by 0.33 points, a statistically significant but small margin obtained with about eight times longer training, and PSO-tuned XGBoost was slightly more accurate and faster at inference. A duplicate-aware split lowered accuracy to 95.26%, showing that conventional evaluation on this benchmark is optimistic by about 1.8 points, and the SSL state, anchor URL and web traffic were the most influential indicators. On an independent dataset the detector reached 98.38% accuracy.

Future work. Future work will replace the discontinued third-party indicators with features that can be computed locally, add visual and textual similarity to known brands to address phishing pages with valid certificates, and evaluate the detector on recent phishing feeds with time-ordered splits.

Real-time applications. The detector can be packaged as a browser extension, an e-mail gateway filter or a campus proxy service that checks each URL before the page is rendered, returning a verdict in about 10 ms and its reasons in about 0.1 s on ordinary hardware.

References

[1] Anti-Phishing Working Group, "Phishing Activity Trends Report, 4th Quarter 2025," APWG, Feb. 2026. [Online]. Available: https://docs.apwg.org/reports/apwg_trends_report_q4_2025.pdf

[2] R. Zieni, L. Massari and M. C. Calzarossa, "Phishing or not phishing? A survey on the detection of phishing websites," *IEEE Access*, vol. 11, pp. 18499–18519, 2023, doi: 10.1109/ACCESS.2023.3247135.

[3] A. Safi and S. Singh, "A systematic literature review on phishing website detection techniques," *Journal of King Saud University – Computer and Information Sciences*, vol. 35, no. 2, pp. 590–611, 2023, doi: 10.1016/j.jksuci.2023.01.004.

[4] A. Zamir, H. U. Khan, T. Iqbal, N. Yousaf, F. Aslam, A. Anjum and M. Hamdani, "Phishing web site detection using diverse machine learning algorithms," *The Electronic Library*, vol. 38, no. 1, pp. 65–80, 2020, doi: 10.1108/EL-05-2019-0118.

[5] M. Al-Sarem, F. Saeed, Z. G. Al-Mekhlafi, B. A. Mohammed, T. Al-Hadhrami, M. T. Alshammari, A. Alreshidi and T. S. Alshammari, "An optimized stacking ensemble model for phishing websites detection," *Electronics*, vol. 10, no. 11, p. 1285, 2021, doi: 10.3390/electronics10111285.

[6] M. S. I. Ovi, M. H. Rahman and M. A. Hossain, "PhishGuard: A multi-layered ensemble model for optimal phishing website detection," arXiv preprint arXiv:2409.19825, 2024, doi: 10.48550/arXiv.2409.19825.

[7] R. M. Mohammad, F. Thabtah and L. McCluskey, "Predicting phishing websites based on self-structuring neural network," *Neural Computing and Applications*, vol. 25, no. 2, pp. 443–458, 2014, doi: 10.1007/s00521-013-1490-z.

[8] R. Mohammad and L. McCluskey, "Phishing Websites," UCI Machine Learning Repository, 2015, doi: 10.24432/C51W2X.

Quantum-Inspired Neural Network for Digital Signature

Aman Kaushik
CSE (Cyber Security) Department
GGSESTC, Bokaro

Kajal Kumari
CSE (Cyber Security) Department
GGSESTC, Bokaro

Avinash Kumar
ME Department
GGSESTC, Bokaro

Arun Prasad Burnwal
BSH Department
GGSESTC, Bokaro

Abstract

Digital signatures play an integral role in cyber security, providing authentication, integrity and non-repudiation. As quantum computing develops, there are increasing concerns that some traditional public key cryptographic algorithms may not be secure in the future. However, neural networks and quantum inspired computational methods offer new opportunities for data analysis and optimization.

This framework uses quantum inspired mathematics, neural networks, cryptographic hashing and a standard post-quantum digital signature algorithm. The quantum inspired neural network performs the data analysis and optimisation whilst the cryptographic algorithm performs signing and verification. Neural networks cannot guarantee the authenticity of a digital signature on their own.

This framework is intended to be analyzed using criteria such as signing time, verification time, memory usage, tamper detection and complexity. As this paper is an intended design and not a full experiment, it will not claim any experimental results in terms of performance and security.

Keywords: Neural Network, Quantum Inspired Neural Network, Digital Signature, Cyber Security, Post-Quantum Cryptography, Cryptographic Hashing, Data Integrity

1. Introduction

Digital signatures are widely used in electronic contracts, banking, e-governance, software distribution, academic records, and other systems where document integrity and signer authenticity are important. A digital-signature scheme generally uses a private key to create a signature and a corresponding public key to verify it. The verification process checks whether the signature corresponds to the signed message and the claimed public key. NIST's Digital Signature Standard describes digital signatures as a mechanism for detecting unauthorized modification and authenticating the signatory [1].

Despite these protections, a valid signature does not always prove that the signing session was safe. A private key may be misused after device compromise, a credential may be stolen, or an attacker may gain access to an authenticated session. In such cases, a cryptographic verifier may correctly report that the signature is valid even though the surrounding activity is suspicious. This creates an opportunity for an intelligent monitoring layer that studies contextual information without interfering with the cryptographic protocol.

This paper proposes a QINN-based risk-scoring framework for this purpose. The model studies patterns such as signing frequency, time of access, device changes, document category, network context, and recent authentication failures. It produces a risk score that can support additional authentication or manual review. The central research question is whether quantum-inspired representations can provide useful feature interactions for this task while remaining efficient on classical computing hardware.

2. Background and Related Work

Quantum-inspired machine learning applies mathematical ideas associated with quantum information to algorithms that execute on classical hardware. Such methods may use amplitude-based representations, phase relationships, probability measurements, or tensor structures. They should not be confused with quantum machine learning that requires a physical quantum processor. A recent weight-constrained neural-network study demonstrates how ideas motivated by amplitude encoding can be expressed through classical neural-network parameters.

Research has also explored quantum and hybrid quantum-classical models for security-related classification tasks. For example, HQRNN-FD applies a hybrid quantum recurrent neural network to financial fraud detection. That work concerns fraud detection rather than digital-signature verification, but it illustrates the broader interest in using quantum based learning architectures to identify suspicious activity. The present proposal differs by using a quantum-inspired model as a classical risk-analysis component around a conventional signature system.

3. Proposed System Architecture

The proposed architecture contains two independent paths: a cryptographic verification path and an intelligent risk-analysis path. Keeping these paths separate is essential because machine learning is probabilistic, whereas cryptographic signature verification must produce a deterministic validity result under the selected algorithm and key.

3.1 Cryptographic Verification Layer

When a signed document is received, the system first applies the normal digital-signature procedure. It checks the document digest, verifies the signature using the signer's public key, validates the certificate or public-key trust information, and checks revocation and validity conditions where applicable. A failed cryptographic check results in rejection. The QINN is not allowed to convert an invalid signature into a valid one, modify the signature, access private keys, or generate cryptographic credentials.

3.2 Context Collection Layer

If the cryptographic checks succeed, the system creates a feature record describing the signing event. Candidate features include: (i) time difference from the signer's normal activity period; (ii) device or browser change; (iii) approximate network or location deviation, where lawful and consented; (iv) number of signing requests within a recent time window; (v) document category or transaction class; (vi) session duration; and (vii) recent failed authentication attempts. Identifiers should be pseudonymised, and raw passwords, private keys, access tokens, and unnecessary document contents must never be included in the model input.

3.3 Quantum-Inspired Neural Network

Let the prepared feature vector be x = [x1, x2, …, xd]. Numerical features are scaled to a consistent range, while categorical features are encoded using a fixed mapping. The vector is then converted into a normalized amplitude-like representation. One possible encoding is:

$$z_j = \sqrt{(r_j / \Sigma\square\ r\square)} \cdot e\text{^}(i\varphi_j), \quad \Sigma_j |z_j|^2 = 1$$

Here, r_j is a non-negative transformation of the j-th input feature and φ_j is a trainable or feature-derived phase. A small positive constant can be added during normalization to avoid division by zero. When input features can be

negative, they are mapped through separate positive and negative channels or transformed before normalization. The encoding is a mathematical representation implemented in ordinary numerical software; it does not create a physical quantum state.

The encoded vector is passed through one or more trainable complex-valued mixing layers. A transformation matrix $U\theta$ may be constrained to satisfy $U\theta^{\dagger}U\theta = I$, where $U\theta^{\dagger}$ is its conjugate transpose. This constraint mirrors the mathematical property of a unitary transformation. After mixing, the model calculates probability-like outputs using $p_j = |(U\theta z)_j|^2$. Because complex amplitudes are combined before their magnitudes are calculated, the resulting values can represent interaction effects between input features. A small classical output layer maps these values to a risk score:

$$\mathbf{R(x) = \sigma(w^{T}p + b), \quad 0 \leq R(x) \leq 1}$$

In this expression, w and b are trainable parameters and σ denotes the sigmoid function. A score near zero represents activity similar to the training examples labelled legitimate, while a score near one indicates a pattern associated with suspicious activity. The score is a risk estimate, not proof of fraud or proof of identity.

3.4 Decision and Response Layer

The response policy uses both the cryptographic result and the learned score. An invalid signature is rejected immediately. A valid signature with a low risk score proceeds normally. A valid signature with an intermediate score may require a one-time password, hardware-key confirmation, or another approved step-up check. A high-risk event may be temporarily held for authorised review. Thresholds must be selected through validation and adjusted to the operational cost of false acceptance and false rejection.

Proposed Methodology

4.1 Dataset Preparation

A practical study would require a labelled dataset of signing events. Each record would contain the permitted contextual features and a target label indicating whether the event was confirmed legitimate or associated with a security incident. Labels should be obtained from verified incident investigations or carefully controlled simulations rather than assumptions based only on unusual behaviour. Because genuine attacks are comparatively rare, the dataset is likely to be imbalanced. Class weighting, carefully designed resampling, or focal loss may be evaluated without allowing synthetic records to leak into the test set.

Data should be divided by time and, where possible, by signer or account. A temporal test set helps measure performance on future activity, while signer-disjoint testing can reveal whether the model merely memorizes individual users. Preprocessing parameters must be fitted on the training set only. Sensitive identifiers should be pseudonymised, and retention periods and access permissions should be defined before collection.

4.2 Training Procedure

The proposed network is trained using labelled examples. The model parameters include the phase values, the parameters of the constrained mixing transformations, and the weights of the output layer. Binary cross-entropy is a suitable initial loss function:

$$\mathbf{L = -[y \log R(x) + (1-y) \log(1-R(x))]}$$

where y is the observed class label. Parameters can be optimized using gradient-based methods. Regularization, dropout, early stopping, and calibration on a separate validation set should be considered to reduce overfitting. The architecture should start small because a complex network is not automatically more secure or more accurate.

4.3 Evaluation Plan

The proposed QINN should be compared with logistic regression, a conventional multilayer perceptron, random forest, and gradient-boosted trees. All models must use the same data splits and feature set. Evaluation should include precision, recall, F1-score, area under the precision-recall curve (PR-AUC), false-positive rate, and inference time. Accuracy alone is insufficient because a system can obtain high accuracy by predicting the majority class. Probability calibration should also be examined so that a risk score of 0.8 has a meaningful operational interpretation.

Ablation tests should remove the phase component, the constrained mixing layer, and selected feature groups one at a time. This would help determine whether the quantum-inspired representation contributes useful information beyond an ordinary neural network. The study should report model size, training time, inference latency, and memory use. No accuracy or efficiency gain can be claimed until these experiments are completed.

5. Security and Deployment Considerations

The QINN is an auxiliary detection mechanism and must not be treated as a cryptographic primitive. Its parameters may be attacked, its training data may contain poisoned records, and an adversary may try to imitate ordinary signing behaviour. Model updates should therefore be authenticated, logged, access-controlled, and tested before deployment. Monitoring for distribution drift is necessary because legitimate user behaviour can change over time.

The system should also be designed to avoid unfair or excessive monitoring. Location, device, and behavioural data can be sensitive; collection should be limited to what is necessary, protected in transit and at rest, and governed by appropriate consent, retention, and audit policies. A risk score should not be the sole reason for permanently denying a person access to a service. Explanations such as "new device and unusual signing frequency" are more useful to security staff than an unexplained probability alone.

Finally, quantum-inspired machine learning does not make a signature quantum-safe. Resistance to attacks from large-scale quantum computers is a property of the cryptographic algorithm, not of the monitoring neural network. NIST has standardized ML-DSA in FIPS 204 and SLH-DSA in FIPS 205 as post-quantum digital-signature mechanisms [2], [3]. A deployment could use one of these algorithms for the cryptographic layer while retaining the proposed QINN as a separate risk-monitoring component.

Limitations and Future Scope

The proposed framework is conceptual and has not been validated through an implementation or benchmark in this paper. Its effectiveness depends on the availability of reliable incident labels, representative user activity, and carefully controlled data collection. Unusual behaviour does not always indicate an attack; legitimate travel, a new device, or a change in work schedule can produce false alarms. Conversely, a compromised account may be used in a way that closely resembles normal activity.

The quantum-inspired representation also introduces design choices that require empirical study. Complex-valued operations and constrained transformations may increase computation compared with a small conventional classifier. Amplitude normalization can reduce the visibility of absolute feature magnitudes unless those values are retained through additional channels. The model may also be difficult to interpret, and no general quantum advantage should be assumed when all computation is performed classically. Future work should implement the architecture, publish reproducible data-preparation procedures, compare it with strong classical baselines, conduct adversarial testing, and evaluate privacy-preserving personalization.

7. Conclusion

This paper presented a proposed quantum-inspired neural network framework for risk-aware digital-signature workflows. The main idea is to combine conventional cryptographic verification with a separate machine-learning layer that evaluates the context in which a signature is produced. Amplitude-like encoding, phase parameters, and interference-inspired transformations provide a possible way to represent relationships among signing-event features using classical computing resources. The model produces a risk estimate that can trigger additional authentication or review, but it cannot replace public-key verification or alter the validity of a cryptographic signature.

The framework therefore addresses a complementary security problem: detecting suspicious use of otherwise valid signing credentials. Its practical value remains an open research question. Controlled experiments, temporal evaluation, comparison with classical models, privacy analysis, and adversarial testing are required before deployment. Separately, post-quantum algorithms such as ML-DSA and SLH-DSA should be selected when quantum-resistant cryptographic protection is required. Combining sound cryptographic standards with carefully evaluated behavioural risk analysis may provide a more layered approach to digital-signature security.

References

1. Gottesman, D., & Chuang, I. Quantum Digital Signatures (2001).
2. Clarke, P. J. et al. Experimental Demonstration of Quantum Digital Signatures Using Phase-Encoded Coherent States of Light.
3. Wallden, P. et al. Quantum Digital Signatures with Quantum Key Distribution Components.
4. Duan, H. A Research on Quantum Digital Signatures.
5. Thomas, M. et al. Comprehensive Review of Post-Quantum Digital Signatures.
6. Aman Kaushik, Avinash Kumar, Arun Prasad Burnwal (2025) "Importance and Need of Digital Signature". In Proceedings of ISTE Faculty Convention 2025 (Bihar & Jharkhand Section), 4-7. ISBN No. – 978-81-956864-6-9.
7. Pratap Kumar, Avinash Kumar, Arun Prasad Burnwal (2025) "Need and Use of Supervised Learning in Cybersecurity". In Proceedings of ISTE Faculty Convention 2025 (Bihar & Jharkhand Section), 8-10. ISBN No. – 978-81-956864-6-9.
8. Aman Kaushik, Avinash Kumar, Arun Prasad Burnwal. "Importance of Mathematics in Digital Signature". Presented In Proceedings of 8th International Conference on Multidisciplinary Perspective on Sustainable Development in Science and Technology 2026, Organized by Maa Shakambari Trust, Greater Noida UP India.
9. Pratap Kumar, Avinash Kumar, Arun Prasad Burnwal. "The Need and Use of Unsupervised Learning in Cybersecurity". Presented In Proceedings of 8th International Conference on Multidisciplinary Perspective on Sustainable Development in Science and Technology 2026, Organized by Maa Shakambari Trust, Greater Noida UP India.

10. Aman Kaushik, Avinash Kumar, Arun Prasad Burnwal. "Quantum Digital Signatures: A Comprehensive Review". Presented in Proceedings of Global Conclave on Sustainable Agriculture, Environmental Health & Bio-Pharmaceutical Sciences (GCSAEHBPS-2026), SARATES\GCSAEHBPS-2026\COP-OP\029.

Quantifying Resampling-Induced Data Leakage in Imbalanced Credit Card Fraud Detection: A Comparative Study of Sampling Strategies and Classification Models

Ashutosh Prasad, Nitesh Kumar, Smita Kishore*
Guru Gobind Singh Educational Society's Technical Campus, Bokaro, *Jharkhand - 827013*
*Corresponding author
Smita Kishore, Assistant Professor, Department of CSE, GGSESTC, Kandra, Bokaro, Jharkhand, e-mail: smitakumari111@gmail.com

ABSTRACT:
Credit card fraud detection is an imbalanced classification problem, and sampling methods such as SMOTE, ADASYN, and SMOTE-Tomek are often used to improve minority-class learning. A common methodological mistake is to apply resampling before the train-test split, which introduces information from future observations into the training process and can inflate measured performance. This study compares a deliberately flawed pre-split resampling procedure with a leakage-safe procedure that applies resampling only inside the training data. The experiment ran on the Kaggle Credit Card Fraud Detection dataset (284,807 transactions, 492 frauds, 0.17% prevalence) across Logistic Regression, Random Forest, and XGBoost models with SMOTE, ADASYN, and SMOTE-Tomek. Results show that the flawed pre-split condition produces dramatically inflated metrics for all models: average precision (mean across samplers) fell from 0.9975 to 0.7265 for Logistic Regression, from 1.0000 to 0.8715 for Random Forest, and from 0.9998 to 0.8272 for XGBoost when resampling was confined to the training data. The distortion was largest for Logistic Regression, smallest for Random Forest, and consistent across all three samplers. Because the flawed condition also shifts the test-set fraud prevalence from 0.17% to 50%, the gap reflects both leakage and a change in the evaluation population. Precision-recall metrics communicate performance more faithfully than accuracy in this rare-event setting. These findings indicate that resampling before train-test separation is a serious validity threat in credit-card fraud detection and should not be used in published evaluation.

Keywords: credit card fraud detection, imbalanced classification, resampling leakage, SMOTE, ADASYN, SMOTE-Tomek, precision-recall evaluation

1. INTRODUCTION:

Credit card fraud detection is one of the clearest examples of a rare event learning problem. In a dataset where non-fraudulent transactions dominate, a model can achieve high accuracy by almost always predicting the majority class. This is not a useful fraud detector. The practical goal is therefore to detect a rare fraud event while keeping false positives under control. This is why precision recall metrics, calibration and threshold selection are so important in this area [2]. Researchers often use oversampling methods such as SMOTE, ADASYN and SMOTE-Tomek to rebalance the data before adjusting a classifier. These methods are valuable when used correctly, but they can also be easily abused. When a sampler is adjusted to the entire data set before train test splitting, the minority samples created by interpolation can draw on information belonging to the future evaluation set. This is a form of data leakage and can lead to optimistic performance estimates that do not generalize to newly observed transactions [1, 3]. The present study asks a targeted methodological question: How does applying resampling before, rather than after, train-test separation affect the measured performance of common classifiers in a fraud dataset? The answer is not accepted. The study treats leakage inflation as a hypothesis to be tested, not an outcome that must occur. It compares an intentionally defective condition with a leak-proof condition and evaluates the gap between three standard models and three common samplers. The goal is not to develop a new algorithm, but to assess the impact of a common methodological error in an area where the risks are real and the consequences of excessive optimism are serious. The study uses the public Kaggle Credit Card Fraud Detection dataset [9], which is highly imbalanced, contains anonymized variables, and only covers a two-day collection period. It is therefore considered an important but limited measure rather than a universal representation of current banking behaviour.

2. LITERATURE REVIEW:

The specialist literature on the topic of "machine learning leaks" goes beyond fraud detection, but is directly relevant here. Kapoor and Narayanan [1] show that leakage is a major factor in irreproducible or misleading scientific results in machine learning-based research. Their argument is relevant to credit card fraud detection because this area relies heavily on benchmark datasets, rare event metrics, and minority class resampling. When a model is evaluated using synthetic examples or data that has already influenced the training pipeline, the score reported is not an estimate of performance on truly unknown transactions. The classic literature on disequilibrium learning is also important. He and Garcia [2] address the challenge of learning from imbalanced data, emphasizing that imbalance is no small annoyance; It changes the objective function, the evaluation threshold and the practical meaning of performance metrics. Their review explains why precision is a poor diagnostic in fraud detection and why metrics like recall, precision, and average precision are often more meaningful. SMOTE is one of the most broadly used oversampling methods. Chawla et al. [3] introduced it to generate synthetic minority examples through interpolation between observations of minorities near neighbors. The method has been widely adopted, but is not immune to abuse. The SMOTE paper itself does not resolve the question of whether the sampler should be applied before or after splitting. Later work, including Batista et al. [4] compared balancing strategies and showed that resampling decisions can significantly change classification behavior. Blagus and Lusa [5] also showed that SMOTE performance can vary in high-dimensional or highly imbalanced environments, making domain-specific evaluation particularly important. The study also draws on modern tabular learning work by Grinsztajn et al. [6], which show that tree-based models remain particularly strong with structured tabular data. This is a methodological reason to include Random Forest and XGBoost instead of relying solely on deep learning or linear baselines. The present study therefore uses a small but strategically selected set of models: Logistic Regression as a transparent baseline, Random Forest as a strong tabular learner and XGBoost as a boosted tree benchmark. The literature does not point to a single accepted estimate of the leakage effect in this data set. Instead, the general principles are established: resampling can alter class balance, oversampling can lead to synthetic observations, and leakage can increase performance if evaluation is not carefully separated. This is the gap that the study closes.

3. DATASET AND PROBLEM SETTING:

The dataset used in this study is the public Kaggle Credit Card Fraud Detection dataset [9]. Its key properties are summarized in Table 1; the counts were verified directly from the downloaded data file. The feature set includes 28

anonymized numerical variables (V1–V28), as well as Time and Amount, and the target is Class, with 1 indicating fraud.

Table 1: Summary of the Credit Card Fraud Detection dataset

Item	Value
Source	Kaggle Credit Card Fraud Detection (mlg-ulb/creditcardfraud)
Total transactions	284,807
Fraudulent transactions	492
Legitimate transactions	284,315
Fraud prevalence	0.001727 (0.1727%)
Number of input features	30 (V1–V28, Time, Amount)
Missing values	0
Collection window	Short two-day period
Notes	Features are anonymized

The dataset is not a perfect representation of modern banking activity. It spans only a short two-day collection window, the features are anonymized, and the context is not necessarily representative of current transaction behavior or current fraud patterns. These are important limitations because they constrain how broadly the conclusions can be generalized. The dataset age and the transformation of the raw financial behavior into anonymized features make the study a useful benchmark for methodology rather than a definitive statement about a live production fraud system.

4. METHODOLOGY:

The study follows a two-condition experimental design. All 18 model-sampler-condition combinations (3 models × 3 samplers × 2 conditions) were executed successfully with a fixed random seed of 42 and an 80/20 stratified train-test split. No runs failed.

a) Hypotheses:

H1: The flawed pre-split resampling condition will produce larger apparent performance estimates than the leakage-safe condition, but the size of the inflation will vary by sampler and model.
H2: Tree-based models may show different sensitivity to resampling leakage than linear models because of their split-based inductive bias and class-distribution sensitivity.
H3: Precision-recall metrics will communicate performance more faithfully than accuracy in a rare-event setting.

b) Models:

Three classifiers were evaluated. Logistic Regression (max_iter=2000, class_weight="balanced"), preceded by standard feature scaling, served as a simple baseline. Random Forest (100 estimators, class_weight="balanced", min_samples_leaf=2) was used for its strong performance on tabular data [6]. XGBoost [8] (100 estimators, max_depth=6, learning_rate=0.05, subsample=0.8, colsample_bytree=0.8, tree_method="hist", device="cuda") served as the boosted-tree benchmark and was run with GPU acceleration. The models were selected because they are common and computationally feasible. All experiments used a fixed random seed of 42.

c) Resampling Methods:

The study compares three oversampling strategies: SMOTE, ADASYN, and SMOTE-Tomek. The samplers were implemented through imbalanced-learn (version 0.14.2) [7] with default parameters and a random seed of 42. Condition B uses imblearn.pipeline.Pipeline because scikit-learn's Pipeline rejects samplers, which implement fit_resample rather than fit_transform. Condition A applies the sampler to the full dataset before splitting, as a deliberately flawed diagnostic.

d) Experimental Conditions:

Condition A is a deliberately flawed pre-split resampling procedure. The sampler is applied to all observations, the resampled data are then split into training and test sets, the classifier is fitted, and the result is evaluated on the resampled test set. It is included for diagnostic comparison only and is invalid for model evaluation, because it uses information from the future evaluation set to create synthetic observations before the split. The purpose is to show how the evaluation distribution can be distorted and how apparently strong scores can arise from contamination.

Condition B is the leakage-safe procedure. The original dataset is split first and the test set is kept untouched; only the training data are resampled, inside a pipeline so that all preprocessing and resampling are fit exclusively on training data. The classifier is fitted on the resampled training set and evaluated on the clean test data. This is the valid baseline for the study and the one that should be interpreted as the primary estimate of generalization performance. A stronger design, in which a final untouched holdout set is separated before any resampling or tuning, was not implemented; this limitation is discussed in Section 7.

e) Data Splitting and Environment:

A stratified 80/20 train-test split with a fixed random seed was used. Stratification preserves the class imbalance in both partitions and is appropriate for a binary fraud classification problem. Cross-validation and repeated seeds were not used, so the results come from a single split. Experiments were run in a Kaggle notebook (Python 3.13, Linux) with scikit-learn 1.6.1, imbalanced-learn 0.14.2, XGBoost 3.4.1, pandas 2.1.3, and NumPy 2.1.3.

f) Evaluation Metrics:

Because fraud is rare, the study relies on more than accuracy. It computes precision, recall, F1-score, average precision (AP), ROC-AUC (supplementary), the confusion-matrix counts (including false positives and false negatives), and the fraud prevalence of each evaluation set. Class predictions used a probability threshold of 0.5. Precision-recall metrics are especially informative because the positive class is rare and because a model's ability to detect fraud is more important than its ability to classify the majority class correctly. Average precision is not the same as integrating a precision-recall curve with the trapezoidal rule; it is the weighted mean of the precisions achieved at each threshold, with the increase in recall from the previous threshold as the weight:

$$AP = \Sigma_n (R_n - R_{n-1}) P_n$$

5. RESULTS:

The experiment ran all 18 model-sampler-condition combinations on the Kaggle Credit Card Fraud Detection dataset. All runs completed successfully; no fits failed. A fixed random seed of 42 was used for all splits and samplers, and the train-test split was 80/20 with stratification. Table 2 reports the metrics for every combination; the raw run-level results and aggregate summaries were exported to results/run_level_results.csv and results/aggregate_results.csv.

a) Main Performance Results:

Table 2: Performance of each model-sampler combination under Condition A (flawed pre-split resampling) and Condition B (leakage-safe training-only resampling)

Model	Sampler	Condition	Precision	Recall	F1	AP	ROC-AUC
Logistic Regression	**SMOTE**	A: pre-split (flawed)	0.9907	0.9702	0.9803	0.9977	0.9974
		B: training-only (safe)	0.0578	0.9184	0.1088	0.7245	0.9708
	ADASYN	A: pre-split (flawed)	0.9903	0.9668	0.9784	0.9972	0.9970
		B: training-only (safe)	0.0187	0.9184	0.0366	0.7304	0.9734

Model	Sampler	Condition	Precision	Recall	F1	AP	ROC-AUC
	SMOTE-Tomek	A: pre-split (flawed)	0.9904	0.9680	0.9791	0.9977	0.9974
		B: training-only (safe)	0.0578	0.9184	0.1088	0.7245	0.9708
Random Forest	**SMOTE**	A: pre-split (flawed)	0.9996	1.0000	0.9998	1.0000	1.0000
		B: training-only (safe)	0.7961	0.8367	0.8159	0.8733	0.9684
	ADASYN	A: pre-split (flawed)	0.9996	1.0000	0.9998	1.0000	1.0000
		B: training-only (safe)	0.8218	0.8469	0.8342	0.8705	0.9678
	SMOTE-Tomek	A: pre-split (flawed)	0.9997	1.0000	0.9998	1.0000	1.0000
		B: training-only (safe)	0.8283	0.8367	0.8325	0.8707	0.9631
XGBoost	**SMOTE**	A: pre-split (flawed)	0.9962	0.9941	0.9951	0.9998	0.9998
		B: training-only (safe)	0.2780	0.8878	0.4234	0.8225	0.9804
	ADASYN	A: pre-split (flawed)	0.9956	0.9932	0.9944	0.9998	0.9998
		B: training-only (safe)	0.2767	0.8980	0.4231	0.8422	0.9793
	SMOTE-Tomek	A: pre-split (flawed)	0.9967	0.9928	0.9948	0.9998	0.9998
		B: training-only (safe)	0.2852	0.8878	0.4318	0.8169	0.9811

Note: AP = average precision. Test-set fraud prevalence was ≈0.50 under Condition A and 0.0017 under Condition B. Condition A scores are not valid estimates of real-world generalization.

In Condition A, the test-set prevalence is approximately 0.50 because the resampled data are balanced before splitting; in Condition B, the test-set prevalence is 0.0017, matching the natural fraud rate of the original dataset. This difference in evaluation population is a confound when comparing conditions directly, as discussed in Section 7.

b) Key Observations:

Condition A (flawed pre-split resampling) produces near-perfect metrics across all models and samplers: average precision ranges from 0.9972 to 1.0000, F1 from 0.9784 to 0.9998, and ROC-AUC from 0.9970 to 1.0000. These scores are not estimates of generalization to unseen, naturally distributed transactions; they reflect evaluation on a resampled test set where synthetic minority observations can appear on both sides of the split.

Condition B (leakage-safe training-only resampling) produces realistic metrics at natural prevalence: average precision ranges from 0.7245 to 0.8733, F1 from 0.0366 to 0.8342, and ROC-AUC from 0.9631 to 0.9811. These are the valid estimates of generalization performance.

The inflation gap between conditions is largest for Logistic Regression (average precision drops from ≈0.9975 to ≈0.7265) and Random Forest shows the smallest drop (from ≈1.0000 to ≈0.8715), with XGBoost in between (from

≈0.9998 to ≈0.8272). The pattern is consistent across all three samplers, indicating that the effect is driven primarily by the evaluation-procedure flaw rather than by any particular sampler.

6. DISCUSSION:

The results confirm that applying resampling before tensile test separation dramatically increases the measured performance for all three models and all three samplers. In Condition A, the average precision for each model-sampler combination is 0. 9972 or greater, with Random Forest essentially reaching 1. 0000. In condition B, the average precision drops to a range of 0. 7245–0. 8733 with a natural cheating prevalence of 0. 17%. This is not a slight overestimation; It is a complete distortion of what the model can actually achieve on invisible transactions. H1 predicted that the faulty pre-split condition would result in larger apparent power estimates than the leak-proof condition, with the level of inflation varying across samplers and models. The results clearly support H1. Any value of condition A far exceeds the value of condition B, and the gap varies depending on the model: the average precision of logistic regression drops from ≈0. 9975 to ≈0. 7265 (a gap of ≈0. 27), random forest from ≈1. 0000 to ≈0. 8715 (≈0. 13), and XGBoost from ≈0. 9998 ≈0. 8272 (≈0. 17). The pattern is consistent across SMOTE, ADASYN, and SMOTE-Tomek, so the bias is due to the error in the scoring process rather than a specific sampler. H2 predicted that tree-based models may have different sensitivity to resampling leakage than linear models. The results partially support this. Logistic regression shows the largest absolute decrease in average precision (≈0. 27), while Random Forest and XGBoost show smaller decreases (≈0. 13 and ≈0. 17, respectively). However, this difference may partly reflect the different baseline performance levels in Condition B and is not a fundamental difference in leak sensitivity. Random Forest achieves the highest average precision under safe conditions (up to 0. 8733), while logistic regression achieves the lowest (up to 0. 7245). Due to the leakage, all models achieve almost perfect results, regardless of the architecture. H3 predicted that precision-recall metrics communicate performance on rare events more reliably than precision. The results support this interpretation. A classifier that classifies every transaction as legitimate would achieve an accuracy of about 99. 83% (1 − 0. 001727) while not detecting fraud at all, so accuracy at this frequency cannot distinguish useful models from useless ones. In contrast, the average precision clearly separates the models in condition B: 0. 7245–0. 7304 for logistic regression, 0. 8705–0. 8733 for random forest, and 0. 8169–0. 8422 for XGBoost. The biased comparison mentioned in the methodology applies here: condition A changes the test set prevalence to ≈50% because the data is balanced before splitting, while condition B is evaluated with the natural prevalence of 0. 17%. The difference in ratings is therefore not a pure estimate of leakage bias alone; it is also a change in the assessment population. However, the direction and size of the gap shows that recalculation before splitting produces misleadingly optimistic results, even though the present design cannot split this gap between leakage and the change in prevalence. The results should be interpreted in light of the limitations of the data set. Due to the short recording period of two days and the anonymized characteristics, these results are more useful for methodological understanding than for making definitive statements about current fraud behavior in production systems..

7. THREATS TO VALIDITY AND LIMITATIONS:

This study has several important limitations. First, the data set is old and limited to a short two-day window, so external validity is limited. Second, the resampling comparison is biased because the faulty condition changes the score distribution by inserting synthetic examples into the test set. This means that a difference in AP or ROC AUC is not a pure estimate of leakage bias; it is also a change in the assessment population. Third, the selection of hyperparameters affects performance, and a small number of default settings are easier to reproduce, but may not be optimal. Fourth, the study examines a small number of models and samplers; This is not a one-size-fits-all comparison of all possible fraud detection pipelines. Fifth, the dataset does not contain detailed transaction metadata, so the results cannot be generalized to all fraud contexts..

8. CONCLUSIONS:

The aim of this study was to answer the question of the extent to which resampling error can distort evaluating fraud detection. The experiment ran all 18 model-sample condition combinations on the Kaggle Credit Card Fraud

Detection dataset (284,807 transactions, 492 fraud cases, 0. 17% prevalence) and provided a clear answer: applying resampling before train-test separation dramatically increases measured performance. In the faulty pre-split condition, the average precision across all models and samplers reached 0. 9972–1. 0000, with Random Forest essentially at 1. 0000. In the leak-proof condition, the average natural prevalence precision fell to 0. 7245–0. 8733. Inflation was greatest in logistic regression (≈0. 27 drop in average precision) and smallest in random forest (≈0. 13 drop), with XGBoost in between (≈0. 17 drop). The pattern was consistent across SMOTE, ADASYN, and SMOTE-Tomek, suggesting that the bias was due to error in the scoring procedure rather than a specific sampler. The contribution of the study is methodological and comparative. No new algorithm is proposed; It quantifies how a common scoring error changes the measured performance of a standard rare event data set. The results confirm that recalculation before train-test separation poses a serious validity threat in credit card fraud detection and should not be used in the published evaluation. Precision-recall metrics, especially average precision, proved to be far more meaningful than accuracy: a trivial classifier that classifies every transaction as legitimate would already achieve an accuracy of around 99. 83%, completely hiding the differences between models. The study has limitations that future work should consider. The comparison between conditions is biased because pre-split resampling changes the test set distribution to a prevalence of ≈50%, while the safe condition is 0. 17%. A stronger design would use a final, untouched holdout set that is separated before resampling. The data set is old (two-day collection window, anonymized characteristics), which limits external validity. Only a train-test split with a single seed was used; Repeated seeds and cross-validation would increase the stability of the estimates. Next steps include adding a clean final holdout set, repeated seeds for stability, temporal validation on a longer data set (if available), and comparison with other imbalance-aware methods such as cost-sensitive learning and threshold adjustment..

REFERENCES:

[1] S. Kapoor and V. Narayanan, "Leakage and the reproducibility crisis in machine-learning-based science," Patterns, vol. 4, no. 9, Sep. 2023, doi: 10.1016/j.patter.2023.100804.

[2] H. He and E. A. Garcia, "Learning from imbalanced data," IEEE Transactions on Knowledge and Data Engineering, vol. 21, no. 9, pp. 1263–1284, Sep. 2009, doi: 10.1109/TKDE.2008.239.

[3] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," Journal of Artificial Intelligence Research, vol. 16, pp. 321–357, Jun. 2002, doi: 10.1613/jair.953.

[4] G. E. A. P. A. Batista, R. C. Prati, and M. C. Monard, "A study of the behavior of several methods for balancing machine learning training data," ACM SIGKDD Explorations Newsletter, vol. 6, no. 1, pp. 20–29, Jun. 2004, doi: 10.1145/1007730.1007735.

[5] R. Blagus and L. Lusa, "SMOTE for high-dimensional class-imbalanced data," BMC Bioinformatics, vol. 14, no. 1, 2013, doi: 10.1186/1471-2105-14-106.

[6] L. Grinsztajn, E. Oyallon, and G. Varoquaux, "Why do tree-based models still outperform deep learning on typical tabular data?" Advances in Neural Information Processing Systems 35, 2022, doi: 10.52202/068431-0037.

[7] imbalanced-learn developers, "Pipeline and resampling examples," official documentation, https://imbalanced-learn.org/stable/.

[8] XGBoost developers, "XGBoost documentation," official documentation, https://xgboost.readthedocs.io/.

[9] Machine Learning Group – ULB, "Credit Card Fraud Detection," Kaggle dataset, https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud.

WISEBUD: "An Offline-First AI-Powered Learning and Student Collaboration Platform"

Shreya Gupta [1], Shreya Anand[2], Aakanksha Sinha[3*]

[1]Student, Department of Computer Science and Engineering, **Government Women's Polytechnic, Bokaro,** Jharkhand
[2]Student, Department of Computer Science and Engineering, **Government Women's Polytechnic, Bokaro,** Jharkhand
[3]Student, Department of Computer Science and Engineering, **Government Women's Polytechnic, Bokaro,** Jharkhand

***Email of the corresponding author: (Shreya Gupta) mk0971747@gmail.com**

Abstract

Access to quality learning support is not always equal. Students may face unreliable internet connectivity, limited learning resources, difficulty getting timely explanations, and a lack of opportunities to learn with peers. WiseBud is proposed as an offline-first, AI-powered student learning and collaboration platform designed to reduce these barriers. Its central idea is to keep essential study support and selected learning resources available even when a student is temporarily disconnected from the internet. The proposed platform brings together study guidance, an AI-based application guide, saved educational content, student discussions, and skill-sharing opportunities in one student-friendly environment. It also aims to support multiple languages, personalised appearance settings, and optional, consent-based nearby student discovery. WiseBud is currently at the prototype and planning stage; several advanced capabilities described in this article are proposed features for future development rather than completed services. This article explains the problem, proposed solution, key features, technical approach, expected benefits, limitations, and future scope of WiseBud.

Keywords: Offline-first learning, artificial intelligence, student collaboration, digital education, peer learning, accessibility.

1. Introduction

Digital learning has created new opportunities for students to access explanations, practice material, educational videos, and academic communities. However, many digital platforms depend on a stable internet connection. When connectivity is slow, expensive, or unavailable, students can lose access to resources at the moment they need them. Learners may also use several separate applications for notes, doubt-solving, communication, and planning, which can make the learning process fragmented.

WiseBud is a proposed student-focused platform that places offline access and collaboration at the centre of its design. Rather than treating internet access as a permanent requirement, the platform is planned to store selected resources and support certain learning activities locally on a device. When connectivity is available, users may update content or access online services; when it is not, previously saved resources and supported local functions can remain available. The goal is not to replace teachers or educational institutions, but to provide an additional support system that helps students study, ask questions, and learn from one another.

2. Problem Statement

Students can experience several connected challenges in their academic journey. First, internet-dependent learning tools may become difficult to use during network interruptions. Second, learners may not always have someone available to explain a topic or guide them through an unfamiliar application. Third, useful educational material can be spread across different platforms, making it harder to organise and revisit. Finally, students often have useful skills and knowledge but lack a simple, structured way to exchange them with peers.

These challenges are not identical for every learner. A student in a low-connectivity area may need offline study material, while another may mainly need peer discussion or help understanding how to use a digital service.

WiseBud addresses these needs through a single proposed platform with modular features that can be developed and improved over time.

3. Proposed Solution and Objectives

WiseBud proposes an offline-first learning environment that combines academic assistance, saved learning resources, and student collaboration. Its design follows a practical principle: essential content that a student has already saved should remain accessible without requiring a continuous internet connection. Features that depend on remote servers, fresh information, or online communication will still require connectivity unless a suitable local or peer-to-peer method is implemented.

- Improve access to previously downloaded study resources during network interruptions.
- Provide a simple guide to help users understand and navigate the application.
- Encourage peer learning through discussions, group interaction, and skill exchange.
- Support a more inclusive experience through planned multilingual and appearance options.
- Give users control over optional location sharing and their participation in student discovery.

4. Major Features of WiseBud

4.1 Offline Study Support

Offline Study Support is one of the main concepts of WiseBud. It aims to help students access supported learning features and previously saved study material without requiring continuous internet connectivity. This can be useful for students who experience network problems or have limited internet access.

4.2 AI Application Guide

The AI Application Guide is designed to help users understand how to use WiseBud. Students can ask questions about application features, navigation, and available options. It aims to make the platform easier to understand, especially for new users.

4.3 Private AI Chat

Private AI Chat provides a personal conversation space between the user and the AI. Students can ask questions and receive individual assistance without sharing their conversations in the public discussion area.

4.4 Public Chat

Public Chat allows students to post their questions so that multiple users can provide different answers and suggestions. This feature encourages interaction, discussion, and the exchange of ideas among students.

4.5 Group Discussion

Group Discussion is planned to provide a space for students to participate in topic-based conversations. It can help students discuss academic subjects, share opinions, and communicate with other learners.

4.6 Student Connect and Location Sharing

WiseBud proposes a Student Connect feature to help students discover nearby learners. Location sharing will be completely optional and based on user consent. Students may choose to share their location within a range of up to 5 kilometres or decline location sharing. Users who do not want to share their location can still use other platform features.

4.7 Contact and Email Support

A direct Contact and Email Support option is planned to help users communicate with the project team if they face any application-related problems or need assistance.

4.8 Multilingual Support and Themes

WiseBud aims to support four languages to make the platform accessible to students with different language preferences. It also proposes Light Mode, Dark Mode, and additional visual theme options to provide a comfortable and personalised interface.

5. Proposed Technical Approach

The planned prototype direction uses standard web technologies such as HTML, CSS, and JavaScript to create a responsive interface that can work across common screen sizes. An offline-first web application approach may use a Progressive Web App (PWA) structure. Service Workers can help manage cached application files, while browser storage such as IndexedDB can store selected user data and downloaded learning resources. These technologies can make previously saved content available without an active connection.

A complete system may later require a backend, database, authentication, content-management services, and secure communication interfaces. An on-device AI model could be explored for selected offline assistance, subject to device performance and model size. Online services would be used for tasks that need current data or communication with remote users. The design should clearly indicate which functions are available offline and which require an internet connection, rather than promising that every feature works offline.

6. Working Concept

The intended user journey begins when a student opens WiseBud and selects a learning or collaboration activity. During an internet-connected session, the student may sign in if required, browse available resources, and download selected material. The platform can then store supported content locally. Later, when the device is offline, the student can open saved resources and use those functions that have been designed to operate locally. When the connection returns, online-dependent features can become available again and stored information may be synchronised where appropriate.

This working concept depends on careful separation between local and online functions. For example, a saved video can be played offline, but searching a remote catalogue for a new video generally needs connectivity. Similarly, a locally available guide may explain app navigation, while a live public discussion with other students requires a communication connection. This distinction is essential for a reliable and honest offline-first experience.

7. Expected Benefits and Applications

If implemented successfully, WiseBud could provide several benefits. Students may be able to revise saved material during network interruptions, organise selected learning resources in one place, and seek peer support through structured discussions. Multilingual support may make the interface easier to approach for a wider range of learners. Optional nearby discovery could help students find peers for academic interaction, while consent controls would allow users to decide whether to share their location.

The platform may be useful for diploma and school students, college learners, self-learners, and student groups working on projects. It could also support peer mentoring, collaborative preparation, and informal skill exchange. These are expected applications; their actual impact would need to be evaluated through testing with students and feedback from educators.

8. Current Development Status and Limitations

WiseBud is currently an evolving prototype and proposed project, not a fully deployed learning platform. The interface and feature concepts are being planned and developed, while major components such as a complete backend, database, working search system, and integrated video library still require development. The AI tutor, offline model support, student messaging, and location-based discovery also need technical implementation and

testing. Therefore, the features described in this article should be understood as proposed capabilities unless they are specifically demonstrated as working in the current prototype.

There are practical limitations to address. Offline storage is restricted by device capacity and browser policies. Large AI models may not run smoothly on every phone. Downloaded content must be kept updated and used with appropriate permissions. Communication between nearby devices can be technically complex and cannot be assumed to work on every device. Location features require strong privacy safeguards, clear consent, and secure handling of data. These issues will guide future design and testing.

9. Future Scope

Future development may focus on completing the backend and database, implementing secure user accounts, adding a searchable and organised learning library, and integrating permitted educational video content. The team may evaluate lightweight on-device AI models for selected offline guidance, while maintaining online support for tasks that require remote processing. Student discussions, group features, and skill exchange can be developed with moderation and reporting tools to promote respectful interaction.

Further improvements may include accessibility testing, performance optimisation for low-end devices, language-quality review, data backup and synchronisation, and user testing with students from different connectivity environments. Feedback from actual users will help determine which features are most useful and should be prioritised.

10. Conclusion

WiseBud presents an idea for a more accessible and collaborative digital learning experience. By combining offline-first resource access with planned AI guidance, student discussion, and peer skill-sharing, the project aims to address some common difficulties faced by learners who depend on digital tools. Its success will depend on transparent offline and online boundaries, reliable technical implementation, responsible privacy practices, and continuous feedback from students. Although WiseBud is still at the prototype and development stage, its vision is to make learning support more flexible and connected—whether or not a student has continuous internet access.

References:

[1] J. Archibald, "The service worker cookbook," Mozilla Developer Network. [Online]. Available: https://serviceworke.rs/
[2] Google Developers, "Progressive Web Apps." [Online]. Available: https://web.dev/explore/progressive-web-apps
[3] Mozilla Developer Network, "IndexedDB API." [Online]. Available: https://developer.mozilla.org/en-US/docs/Web/API/IndexedDB_API
[4] W. Holmes, M. Bialik, and C. Fadel, Artificial Intelligence in Education: Promises and Implications for Teaching and Learning. Boston, MA, USA: Center for Curriculum Redesign, 2019.
[5] P. Dillenbourg, Collaborative Learning: Cognitive and Computational Approaches. Oxford, U.K.: Elsevier, 1999.
[6] R. C. Clark and R. E. Mayer, E-Learning and the Science of Instruction: Proven Guidelines for Consumers and Designers of Multimedia Learning, 4th ed. Hoboken, NJ, USA: Wiley, 2016.
[7] OWASP Foundation, "OWASP Application Security Verification Standard." [Online]. Available: https://owasp.org/www-project-application-security-verification-standard/
[8] World Wide Web Consortium (W3C), "Web Content Accessibility Guidelines (WCAG) 2.2," 2023. [Online]. Available: https://www.w3.org/TR/WCAG22/
[9] National Institute of Standards and Technology, "NIST Privacy Framework: A Tool for Improving Privacy through Enterprise Risk Management, Version 1.0," Jan. 2020. [Online]. Available: https://www.nist.gov/privacy-framework
[10] UNESCO, Guidance for Generative AI in Education and Research. Paris, France: UNESCO, 2023. [Online]. Available: https://unesdoc.unesco.org/ark:/48223/pf0000386693

Changes In Technical Education Using Artificial Intelligence: A Review of Emerging Practices, Opportunities and Challenges

Ashish Raj[1] and Salim Ahmad[2]*
[1]Student, Department of Civil Engineering, GGSESTC, Kandra, Bokaro
[2*]Assistant Professor, Department of ECE, GGSESTC, Kandra, Bokaro; Email salim.ahmad2008@gmail.com

Abstract

Artificial Intelligence (AI) is reshaping technical education by changing how engineering, computing and vocational subjects are taught, practised and assessed. AI-enabled learning platforms can provide personalised feedback, intelligent tutoring, simulation, predictive support and assistance with coding and design. In laboratories and workshops, AI can support virtual experimentation, fault diagnosis, robotics and data-driven problem solving. These developments can help institutions connect classroom learning with the rapidly changing needs of industry. However, successful adoption requires reliable infrastructure, teacher preparation, curriculum renewal, data protection, academic integrity and equal access. This paper reviews major changes brought by AI in technical education and proposes a human-centred framework for responsible implementation. It uses recent international reports to describe changing skill requirements and argues that AI should strengthen, rather than replace, teachers and practical learning.

Keywords: Artificial intelligence, technical education, engineering education, adaptive learning, vocational training, digital skills.

1. Introduction

Technical education prepares learners for occupations that depend on applied science, engineering, information technology and practical competence. Traditional teaching commonly relies on lectures, textbooks, fixed laboratory schedules and periodic examinations. These methods remain important, but they may not always provide individual support or keep pace with fast-changing technologies. AI introduces systems that can analyse information, recognise patterns, generate content and support decisions. Its growing presence in workplaces means that technical institutions must both use AI as an educational tool and teach students how to work responsibly with AI-based systems.

The World Economic Forum (WEF, 2025) reports that AI and big data, networks and cybersecurity, and technological literacy are among the fastest-growing skill areas expected by employers. It also estimates that 39% of workers' existing skills may be transformed or become outdated during 2025–2030. These trends make curriculum renewal and continuous learning essential.

2. Objectives

The objectives of this paper are: (i) to identify key ways AI is changing technical teaching and learning; (ii) to examine benefits and risks for students, teachers and institutions; and (iii) to suggest practical steps for responsible AI integration.

3. Major Changes in Technical Education Through AI

3.1 Personalised and adaptive learning

AI-enabled learning platforms can analyse a learner's responses, pace and areas of difficulty. They can recommend additional explanations, practice questions or learning resources. In technical subjects, this can help students revise mathematics, programming, electronics and engineering concepts at a suitable level. Teachers can use learning

analytics to identify topics that require additional classroom attention. Such systems should support teacher judgement and should not make high-stakes decisions without human review.

3.2 Smart laboratories and simulation

AI can be combined with virtual laboratories, digital twins and simulation software to let students explore systems before working with physical equipment. Engineering students may simulate circuits, manufacturing processes or mechanical behaviour; computer science students may test algorithms and datasets; and vocational learners may practise troubleshooting in simulated environments. These tools can improve preparation and reduce some equipment constraints, but they cannot fully replace hands-on laboratory work, workshop safety training or real-world measurement.

3.3 AI-assisted programming, design and projects

Generative AI tools can explain code, suggest debugging steps, help document projects and support early-stage design. Used carefully, they can help learners compare solutions and understand complex concepts. Students must still demonstrate their own reasoning, test generated outputs and cite or disclose AI assistance according to institutional rules. Instructors can design assessments that require demonstrations, oral explanations, design decisions and practical verification.

3.4 Intelligent assessment and feedback

AI can assist with formative quizzes, automated feedback, language support and analysis of learning progress. In practical courses, computer vision and sensor data may help identify errors in a procedure or machine operation. Automated assessment must be transparent, accessible and checked for bias. Final evaluation of competence—especially in safety-critical trades—should include direct observation by qualified instructors.

3.5 Curriculum and industry alignment

AI tools can help institutions review job advertisements, occupational standards and industry feedback to identify emerging competencies. The OECD (2026) highlights the potential for AI to support labour-market analysis, competency mapping and the updating of vocational education and training programmes. Curriculum decisions should nevertheless involve employers, teachers, students and professional bodies.

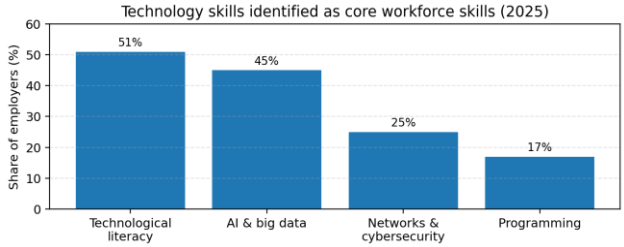


Figure 1. Selected technology skills identified as core workforce skills by surveyed employers. Source: WEF, Future of Jobs Report 2025. Values represent employer responses, not student performance or AI adoption rates.

4. Benefits for Students and Institutions

• Improved learning support: adaptive explanations and immediate feedback can help students address knowledge gaps.
• More opportunities for practice: simulation and virtual environments can supplement limited laboratory access.
• Stronger employability: AI literacy, data analysis, programming and cybersecurity can be integrated into technical programmes.
• Better teaching support: educators may use AI to prepare examples, generate practice material and identify common misconceptions.

• Curriculum responsiveness: analysis of industry trends can help institutions update course content and short-term training.

5. Challenges and Risks

5.1 Accuracy and overdependence

AI-generated answers may be incorrect, incomplete or outdated. Students who copy outputs without understanding them may weaken their problem-solving skills. Institutions should require verification, practical testing and explanation of work.

5.2 Data privacy, security and bias

Learning platforms may process personal information, assessment records or sensitive institutional data. Institutions need clear rules on consent, data minimisation, secure storage and approved tools. AI systems may also reproduce bias in training data, so their outputs should be reviewed for fairness.

5.3 Unequal access and infrastructure

Reliable internet, suitable devices, licensed software and accessible interfaces are not equally available to all learners. If AI-based learning becomes compulsory without support, it may widen existing educational inequalities. Colleges should provide campus access, low-bandwidth alternatives and inclusive learning support.

5.4 Teacher readiness and academic integrity

Teachers need professional development in AI literacy, prompt design, assessment redesign and responsible data use. Academic policies should distinguish acceptable assistance from plagiarism or misrepresentation. Students should disclose substantial AI assistance and remain accountable for submitted work.

6. A Practical Framework for AI Integration

A phased approach is recommended. First, assess institutional needs, infrastructure and student access. Second, select a small number of approved tools for clearly defined learning outcomes. Third, train teachers and students in verification, privacy, bias and academic integrity. Fourth, pilot the tools in selected courses and compare learning outcomes with existing methods. Finally, evaluate accessibility, cost, reliability and student competency before wider deployment.

Area	Suggested action	Evidence of progress
Teaching	Train faculty; use AI for feedback and examples	Teacher training and course review
Laboratories	Combine simulation with supervised practical work	Practical competency assessment
Governance	Approve tools; set privacy and disclosure rules	Policy compliance and audits
Equity	Provide devices, access and alternative formats	Participation and access data

7. Discussion

The evidence suggests that AI is not simply another classroom technology; it affects the knowledge and competencies that technical education must develop. Learners need technical skills such as data handling, programming, cybersecurity and AI fundamentals, alongside communication, teamwork, creativity, ethical judgement and practical problem-solving. WEF (2025) identifies analytical thinking as a leading core skill and reports that employers increasingly value technological literacy and AI-related capabilities. Therefore, technical programmes should integrate AI across disciplines rather than treating it only as a separate subject.

At the same time, the purpose of technical education remains the development of competent people who can design, build, maintain, test and improve real systems. AI can accelerate access to information and provide practice

opportunities, but it cannot take responsibility for professional decisions, workplace safety or ethical conduct. The strongest model is a partnership in which teachers guide learning, students actively construct knowledge, AI provides bounded assistance, and practical assessment confirms competence. UNESCO's 2026 guide to AI integration in technical and vocational education similarly emphasises institutional planning, human agency, equity and safety.

8. Recommendations

i. Introduce AI literacy, data ethics and responsible-use practices across engineering, diploma and vocational curricula.
ii. Train teachers to evaluate AI outputs and design assessments that reward understanding, originality and practical skill.
iii. Expand access to supervised virtual laboratories while preserving hands-on workshops and equipment-based assessment.
iv. Establish institutional policies for privacy, cybersecurity, transparency, bias review and disclosure of AI assistance.
v. Build partnerships with industry to update competencies and provide internships, projects and real-world datasets.
vi. Evaluate AI initiatives regularly using learning outcomes, student feedback, accessibility, cost and employability indicators.

9. Conclusion

Artificial Intelligence is changing technical education through personalised learning, intelligent feedback, simulation, AI-assisted design and more responsive curriculum development. These applications can improve learning opportunities and strengthen alignment with emerging industry requirements. Their benefits, however, depend on reliable infrastructure, prepared teachers, ethical governance and equitable access. AI should be adopted as a supportive tool within a human-centred education system, not as a substitute for teachers, practical experience or independent thinking. Institutions that combine digital innovation with strong foundational knowledge and hands-on competence will be better prepared to educate a future-ready technical workforce.

References

1. UNESCO-UNEVOC. (2026). Integrating AI in TVET: A practical guide for institutions. UNESCO International Centre for Technical and Vocational Education and Training. https://www.unesco.org/en/articles/integrating-ai-tvet-practical-guide-institutions

2. OECD. (2026). Developing Vocational Education and Training with Artificial Intelligence. OECD Publishing. https://doi.org/10.1787/e9f76b4e-en

3. World Economic Forum. (2025). The Future of Jobs Report 2025. World Economic Forum. https://www.weforum.org/publications/the-future-of-jobs-report-2025/

4. World Economic Forum. (2025). Skills Outlook, The Future of Jobs Report 2025. https://www.weforum.org/publications/the-future-of-jobs-report-2025/3-skills-outlook

A Critical Review of Recently Proposed Metaheuristic Algorithms: Taxonomy, Benchmark Performance, and Open Research Gaps

Prince Kumar[1] ,Rohit Verma[2], and Manoj Kumar[3]

[1]Student, Department of EEE, Guru Gobind Singh Educational Society's Technical Campus, Kandra, Bokaro, Jharkhand
[2] Assistant Professor, Department ECE, Guru Gobind Singh Educational Society's Technical Campus, Kandra, Bokaro, Jharkhand
[3]Assistant Professor, Department of Mechanical Engineering, Guru Gobind Singh Educational Society's Technical Campus, Kandra, Bokaro, Jharkhand
*Email of the corresponding author: rhitverm@gmail.com[2] , manoj.mehta68@gmail.com[3]

Abstract

Metaheuristic algorithms (MAs) have become a standard tool for optimization problems that are non-linear, high-dimensional, or otherwise unsuitable for classical mathematical programming. Because hundreds of new MAs have been proposed in little more than a decade, practitioners struggle to judge which ones are worth adopting. This review examines the comparative study by Amini-Valashani and Mirzakuchaki (2025), which classifies MAs into nine families by source of inspiration and benchmarks fifteen algorithms published in 2023 on twenty classical test functions and the ten CEC-2019 "100-Digit Challenge" problems. We summarise the taxonomy, the algorithms studied, the experimental design, and the reported findings on accuracy, stability, computational cost, an accuracy–time figure of merit, and convergence speed. We then critically assess the methodology, noting strengths (breadth, multiple dimensions, a combined efficiency metric) and limitations (default parameters, short iteration budgets, absence of formal statistical tests, and sensitivity of the figure of merit to zero-error results). The review closes with recommendations for more rigorous benchmarking and for future work on MAs.

Keywords: metaheuristic algorithms, global optimization, benchmark functions, CEC-2019, nature-inspired computing, performance evaluation

Introduction

Optimization underpins engineering design, scheduling, energy systems, machine learning, and many other fields. Classical approaches such as linear programming, non-linear programming, and dynamic programming remain valuable, but each has well-known weaknesses: linear programming requires linear objectives, non-linear programming tends to become trapped in local optima on difficult landscapes, and dynamic programming suffers from rapidly growing memory and computation as problem size increases (Amini-Valashani & Mirzakuchaki, 2025). These limitations motivated stochastic, population-based search methods, collectively called metaheuristic algorithms.

MAs are generally inspired by natural, social, physical, or mathematical processes. They do not guarantee the exact optimum, but they often return high-quality solutions within a reasonable time and apply to continuous, discrete, single-objective, and multi-objective problems (Rajwar et al., 2023). The field has grown very quickly: the surveyed paper, citing Rajwar et al. (2023), notes that more than 500 MAs have been proposed, over 350 of them in the last decade. Such growth raises a practical question: which of the many new algorithms genuinely perform well.

This review addresses that question through a close reading of Amini-Valashani and Mirzakuchaki (2025), a comparative study of fifteen MAs introduced in 2023. Our objectives are to (a) summarise the taxonomy and algorithms covered, (b) report the experimental design and principal findings, (c) critically evaluate the methodology and the strength of its conclusions, and (d) outline directions for future research.

Taxonomy of Metaheuristic Algorithms

Amini-Valashani and Mirzakuchaki (2025) group MAs into nine categories according to their source of inspiration. Table 1 summarises these categories with representative algorithms drawn from the paper and its cited literature.

Table 1

Nine Families of Metaheuristic Algorithms and Representative Examples

Family	Inspiration	Representative algorithms
Evolutionary	Natural selection, crossover, mutation	Genetic Algorithm (Koza, 1992); Barnacles Mating Optimizer
Swarm intelligence	Collective animal behaviour	Ant Colony Optimization; Particle Swarm Optimization (Kennedy & Eberhart, 1995)
Natural science	Physical and chemical laws	Simulated Annealing (Kirkpatrick et al., 1983); Equilibrium Optimizer
Mathematics-based	Mathematical functions and operators	Sine Cosine Algorithm (Mirjalili, 2016); Arithmetic Optimization Algorithm
Human-based	Human thinking, teaching, social behaviour	Political Optimizer; Group Teaching Optimization
Sports (game)-based	Rules and play of sports	League Championship Algorithm; Volleyball Premier League
Plant-based	Plant growth and reproduction	Flower Pollination Algorithm; Artificial Plant Optimization
Art-based	Music and design	Harmony Search; Melody Search
Disease and virus-based	Epidemics and viruses	Coronavirus Herd Immunity Optimizer; COVIDOA

The taxonomy is convenient but is based only on the metaphor behind each algorithm. As discussed in the critical discussion below, metaphor alone says little about the mathematical mechanisms that determine performance.

Algorithms Evaluated

The study evaluates fifteen algorithms published in 2023 (Table 2). Seven are swarm-based, two are physics/chemistry-based, two are mathematics-based, two are human-based, one is sports-based, and one is disease/virus-based.

Table 2

Metaheuristic Algorithms Published in 2023 and Evaluated by Amini-Valashani and Mirzakuchaki (2025)

Algorithm (acronym)	Family	Inspiration
Osprey Optimization Algorithm (OOA)	Swarm	Ospreys hunting fish
Coati Optimization Algorithm (COA)	Swarm	Coati foraging and predator escape
American Zebra Optimization Algorithm (AZOA)	Swarm	Social behaviour and leadership of zebras
Dung Beetle Optimizer (DBO)	Swarm	Rolling, dancing, foraging, stealing, breeding of dung beetles
Enhanced Snake Optimizer (ESO)	Swarm	Snake mating process
Tyrannosaurus Optimization Algorithm (TROA)	Swarm	Hunting by T. rex
Genghis Khan Shark Optimizer (GKSO)	Swarm	Hunting, movement, foraging, and colour-change escape of the shark
Chernobyl Disaster Optimizer (CDO)	Physics/chemistry	Reactor explosion and radiation particles
Propagation Search Algorithm (PSA)	Physics/chemistry	Voltage and current waves on transmission lines
Exponential Distribution Optimizer (EDO)	Mathematics	Exponential distribution model
Subtraction-Average-Based Optimizer (SABO)	Mathematics	Subtraction average of search agents
Skill Optimization Algorithm (SOA)	Human	Acquiring and improving skills
Mountaineering Team-Based Optimization (MTBO)	Human	Cooperation of mountaineering teams
Golf Optimization Algorithm (GOA)	Sports	Strategy and player conduct in golf
Corona Virus Search Optimizer (CVSO)	Disease/virus	Spread, search, and death of a virus in societies

Original sources include Dehghani and Trojovský (2023) for OOA, Dehghani et al. (2023) for COA, Xue and Shen (2023) for DBO, Yao et al. (2023) for ESO, Hu et al. (2023) for GKSO, Shehadeh (2023) for CDO, Qais et al. (2023) for PSA, Abdel-Basset et al. (2023) for EDO, and Trojovský and Dehghani (2023) for SABO.

Experimental Design of the Reviewed Study

Benchmark Problems

The authors assembled twenty classical benchmark functions (F01–F20): eight unimodal (e.g., Sphere, Trid, Zakharov, Schwefel 2.21 and 2.22), six variable-dimension multimodal (e.g., Salomon, Alpine 1, Drop-Wave, Levy, Styblinski–Tang, Michalewicz), and six fixed-dimension multimodal (e.g., Shekel's Foxholes, Six-Hump Camel Back, Bukin N.6, Egg Holder, Colville). Variable-dimension functions were run at 10, 20, and 30 dimensions. The second test set consisted of the ten CEC-2019 "100-Digit Challenge" problems, with dimensions of 9, 16, 18, or 10 (Price et al., 2018). The function collection draws on benchmark compilations such as Plevris and Solorzano (2022).

Parameter Settings and Protocol

All algorithms used the default control parameters from their publicly available MATLAB implementations. Population size was set to ten times the problem dimension, and the maximum number of iterations to twenty times the dimension, a modification of an earlier formula that produced negative values for some low-dimensional functions. Each algorithm was run 30 times per problem, on MATLAB R2019a, a Core i5 2.30 GHz CPU, and 8 GB RAM.

Performance Indicators

Five types of analysis were reported:

1. Accuracy: best, worst, and mean objective values, with ranking based on the mean.
2. Stability: the standard deviation across 30 runs.
3. Computational cost: theoretical time complexity and measured average run time.
4. Accuracy–time trade-off: a figure of merit, FOM = 1 / (MAE × T), where MAE is the mean absolute error relative to the known optimum and T is the average run time.
5. Convergence speed: convergence curves from a single representative run (run 15).

Summary of Reported Findings

Accuracy

On the unimodal functions (F01–F08), ESO and GKSO achieved the best mean values, followed by OOA and COA, whereas MTBO and CVSO ranked lowest in most cases. For the variable-dimension multimodal functions, ESO and GKSO again led in most cases, with SOA and GOA performing particularly well on F11 and F12. On the fixed-dimension functions, ESO was best on F15 and F18, while SOA, GKSO, and COA led on F15, F19, and F20, respectively (as stated in the original text). On CEC-2019, GKSO ranked first by mean on CEC04 and CEC08–CEC10; ESO, CVSO, and SOA were the three best on CEC05–CEC07; and many algorithms reached the same value on CEC02 and CEC03.

Stability

GKSO produced the lowest standard deviations on nearly all unimodal functions (F04 being the exception), with ESO, COA, and OOA close behind. Stability leaders varied across the multimodal functions. For the fixed-dimension and CEC-2019 problems, the TROA and CDO algorithms showed the lowest standard deviations in most cases.

Computational Time

CDO and PSA had the lowest theoretical complexity and, in the majority of unimodal cases, the shortest measured run times. DBO and TROA were fastest on several multimodal and fixed-dimension problems, and TROA was fastest on most of the later CEC-2019 problems. The algorithms with the strongest accuracy (ESO, GKSO) were not the fastest in absolute terms.

Accuracy–Time Figure of Merit

Counting the number of problems on which each algorithm obtained the highest FOM, the authors report OOA 11, COA 9, ESO 9, SOA 8, GOA 7, PSA 6, TROA 5, GKSO 5, EDO 4, AZOA 3, DBO 3, CDO 2, MTBO 2, CVSO 1, and SABO 0. Thus OOA, COA, ESO, and SOA offered the best compromise between precision and speed even when they were not the most accurate or the fastest individually.

Convergence

In most of the convergence plots for F01–F20, ESO converged fastest, followed by GKSO, COA, and OOA, while MTBO and CVSO converged slowest. On the CEC-2019 problems convergence was generally slower, which the authors attribute to the greater complexity of those problems. The authors conclude that ESO and GKSO performed best overall, with COA, OOA, and SOA next, and with PSA and CDO preferable when computation time is the main concern.

Critical Discussion

Strengths

- Breadth and currency. The study covers fifteen algorithms from a single publication year across nine metaphor families, offering a useful snapshot of the newest generation of MAs.
- Varied problem types. The use of unimodal, multimodal, fixed-dimension, and CEC-2019 problems at several dimensions gives a more complete picture than single-suite evaluations.
- Combined metric. The FOM acknowledges that accuracy and runtime must be considered together, which is often neglected in algorithm-proposal papers.
- Reproducibility support. The authors list the sources of the MATLAB implementations, enabling replication.

Limitations

1. Default parameters. Using each algorithm's default settings is convenient and arguably fair, but defaults were typically tuned by the original authors for their own benchmark choices, so the comparison may favour some algorithms over others.
2. Small evaluation budgets. The iteration limit of 20 × dimension is modest, especially for 30-dimensional and CEC-2019 problems. Rankings after a short budget may differ from rankings at convergence, and a budget expressed in function evaluations (rather than iterations) would be fairer, since algorithms perform different numbers of evaluations per iteration.
3. No formal statistical testing. Rankings rely on mean values and standard deviations. Non-parametric tests, such as the Wilcoxon rank-sum or Friedman test, are standard practice in the CEC community and would indicate whether observed differences are significant.
4. Heavy reliance on zero-error cases. Many algorithms reach exactly zero (or underflow-level) errors on the simpler unimodal functions, producing ties that make rankings uninformative. The FOM also becomes

unbounded or capped when MAE is zero; the very large and repeated identical FOM entries in the reported table suggest that such cases need special handling, which should be documented.

5. Single-run convergence curves. Convergence behaviour is shown for only one run (run 15), which is anecdotal. Averaged or median curves with variability bands would be more reliable.
6. Hardware-dependent run times. Measured times come from one modest machine and one MATLAB version, and implementation quality varies between authors; timing comparisons should be treated as indicative.
7. Presentation inconsistencies. Some statements in the narrative do not match perfectly with the tables or abstract (for example, the section on fixed-dimension stability refers to F14 although that function belongs to the preceding group, and the abstract highlights SOA while the conclusion lists COA and OOA ahead of it). Readers should consult the raw tables when using the rankings.
8. Metaphor-based classification. The taxonomy is organised around inspiration sources. Critics argue that novel metaphors can disguise algorithms that are mathematically similar to existing ones; a classification based on search operators would be more informative. This point is our observation rather than a claim made in the reviewed paper.
9. Benchmark scope. Classical test functions and CEC-2019 problems do not represent constrained, multi-objective, or real-world engineering problems. By the no-free-lunch principle, the best algorithm on these benchmarks need not be best elsewhere.

Implications for Practitioners

For practitioners with accuracy as the priority, the reviewed evidence supports shortlisting ESO and GKSO, with COA, OOA, and SOA as alternatives. Where computational time is the main restriction, PSA, CDO, and TROA are candidates, while OOA, COA, ESO, and SOA offer a balance. These conclusions should be confirmed on the target problem class before adoption.

Future Research Directions

- Rigorous benchmarking: use equal function-evaluation budgets, statistical tests, parameter tuning for each competitor, and larger dimensionality.
- Wider test suites: include newer CEC suites, constrained engineering design problems, and multi-objective problems.
- Operator-level analysis: analyse which exploration and exploitation mechanisms explain the success of ESO and GKSO.
- Hybridisation: combine fast algorithms (e.g., PSA, CDO) with accurate ones (e.g., ESO, GKSO).
- Adaptive parameter control: reduce reliance on hand-tuned or default parameters.
- Open resources: share code, seeds, and raw results to support replication.

Conclusion

Amini-Valashani and Mirzakuchaki (2025) provide a broad and practically useful comparison of fifteen recent metaheuristic algorithms. Their results indicate that the Enhanced Snake Optimizer and the Genghis Khan Shark Optimizer deliver the best accuracy, stability, and convergence on the tested functions, that the Propagation Search Algorithm and Chernobyl Disaster Optimizer are the fastest in most cases, and that OOA, COA, ESO, and SOA achieve the best accuracy–time balance. At the same time, the use of default parameters, short iteration budgets, single-run convergence plots, and descriptive rather than inferential statistics means the rankings should be treated as indicative rather than definitive. More standardised and statistically grounded benchmarking will be needed to separate genuinely better algorithms from those that merely perform well under a particular experimental set-up.

References:

1. Abdel-Basset, M., El-Shahat, D., Jameel, M., & Abouhawwash, M. (2023). Exponential distribution optimizer (EDO): A novel math-inspired algorithm for global optimization and engineering problems. *Artificial Intelligence Review, 56*(9), 9329–9400. https://doi.org/10.1007/s10462-023-10403-9
2. Amini-Valashani, M., & Mirzakuchaki, S. (2025). Performance evaluation of latest meta-heuristic algorithms in finding optimum value of mathematical functions and problems. *Arabian Journal for Science and Engineering, 50*, 15365–15405. https://doi.org/10.1007/s13369-024-09807-8
3. Dehghani, M., Montazeri, Z., Trojovská, E., & Trojovský, P. (2023). Coati optimization algorithm: A new bio-inspired metaheuristic algorithm for solving optimization problems. *Knowledge-Based Systems, 259*, Article 110011. https://doi.org/10.1016/j.knosys.2022.110011
4. Dehghani, M., & Trojovský, P. (2023). Osprey optimization algorithm: A new bio-inspired metaheuristic algorithm for solving engineering optimization problems. *Frontiers in Mechanical Engineering, 8*, Article 1126450. https://doi.org/10.3389/fmech.2022.1126450
5. Hu, G., Guo, Y., Wei, G., & Abualigah, L. (2023). Genghis Khan shark optimizer: A novel nature-inspired algorithm for engineering optimization. *Advanced Engineering Informatics, 58*, Article 102210. https://doi.org/10.1016/j.aei.2023.102210
6. Kennedy, J., & Eberhart, R. (1995). Particle swarm optimization. In *Proceedings of ICNN'95 – International Conference on Neural Networks* (Vol. 4, pp. 1942–1948). IEEE. https://doi.org/10.1109/ICNN.1995.488968
7. Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. P. (1983). Optimization by simulated annealing. *Science, 220*(4598), 671–680. https://doi.org/10.1126/science.220.4598.671
8. Koza, J. R. (1992). *Genetic programming: On the programming of computers by means of natural selection*. MIT Press.
9. Mirjalili, S. (2016). SCA: A sine cosine algorithm for solving optimization problems. *Knowledge-Based Systems, 96*, 120–133. https://doi.org/10.1016/j.knosys.2015.12.022
10. Plevris, V., & Solorzano, G. (2022). A collection of 30 multidimensional functions for global optimization benchmarking. *Data, 7*(4), Article 46. https://doi.org/10.3390/data7040046
11. Price, K. V., Awad, N. H., Ali, M. Z., & Suganthan, P. N. (2018). *Problem definitions and evaluation criteria for the 100-digit challenge special session and competition on single objective numerical optimization* [Technical report]. Nanyang Technological University.
12. Qais, M. H., Hasanien, H. M., Alghuwainem, S., & Loo, K. H. (2023). Propagation search algorithm: A physics-based optimizer for engineering applications. *Mathematics, 11*(20), Article 4224. https://doi.org/10.3390/math11204224
13. Rajwar, K., Deep, K., & Das, S. (2023). An exhaustive review of the metaheuristic algorithms for search and optimization: Taxonomy, applications, and open challenges. *Artificial Intelligence Review, 56*, 13187–13257. https://doi.org/10.1007/s10462-023-10470-y
14. Shehadeh, H. A. (2023). Chernobyl disaster optimizer (CDO): A novel meta-heuristic method for global optimization. *Neural Computing and Applications, 35*, 10733–10749. https://doi.org/10.1007/s00521-023-08261-1
15. Trojovský, P., & Dehghani, M. (2023). Subtraction-average-based optimizer: A new swarm-inspired metaheuristic algorithm for solving optimization problems. *Biomimetics, 8*(2), Article 149. https://doi.org/10.3390/biomimetics8020149
16. Xue, J., & Shen, B. (2023). Dung beetle optimizer: A new meta-heuristic algorithm for global optimization. *The Journal of Supercomputing, 79*, 7305–7336. https://doi.org/10.1007/s11227-022-04959-6
17. Yao, L., Yuan, P., Tsai, C.-Y., Zhang, T., Lu, Y., & Ding, S. (2023). ESO: An enhanced snake optimizer for real-world engineering problems. *Expert Systems with Applications, 230*, Article 120594. https://doi.org/10.1016/j.eswa.2023.120594

AI-Driven Smart Suraksha Helmet with Edge AI for Real-Time Hazard Detection

Bhola Kumar Mahato
Department of Computer Science Engineering (Data Science)
GGSESTC, Bokaro

Sahin Parween
Department of Computer Science Engineering
GGSESTC, Bokaro

Abhilash Ghosh
Department of Electrical & Electronics Engineering
GGSESTC, Bokaro

ABSTRACT: Jharkhand is a major coal-producing state contributing approximately 19–20% of India's total coal output. Despite regulatory oversight by the Directorate General of Mines Safety (DGMS), gas accumulation and roof/side falls continue to cause fatalities and serious injuries in underground and open-cast operations. Conventional mining helmets provide only passive physical protection and lack real-time sensing or decision-making capability. This paper presents the AI-Driven Smart Suraksha Helmet, a low-cost edge-AI wearable system built around the ESP32 microcontroller, MQ-2 gas sensor, MPU6050 inertial measurement unit, and infrared proximity sensor, with LoRa (SX1278) communication. The helmet performs local inference for gas-trend prediction, fall detection, and helmet-compliance monitoring, generating alerts in under 0.5 seconds without requiring internet connectivity—an essential feature for deep underground workings where cellular networks are unavailable. The prototype is estimated at approximately ₹3,500, adding intelligent safety functions at an incremental cost of roughly ₹2,700 over a conventional helmet. Laboratory testing demonstrated approximately 92% overall accuracy in hazard classification. The design aligns with Atmanirbhar Bharat objectives by utilising readily available components and addresses the practical constraints of Indian coal mines.

Keywords—Smart Helmet, Edge-AI, ESP32, LoRa, Coal Mine Safety, Gas Detection, Fall Detection, IoT, DGMS

I. INTRODUCTION:

Coal mining remains critical to India's energy security. Jharkhand hosts a large number of operational mines under Central Coalfields Limited (CCL), Bharat Coking Coal Limited (BCCL), and private operators. According to recent Ministry of Coal data, the state produced over 200 million tonnes in recent fiscal years, representing roughly one-fifth of national output. Despite improvements in mechanisation and regulatory enforcement, accidents continue to occur. DGMS statistics indicate that ground movement (including roof and side falls) and gas-related incidents remain significant contributors to fatalities in underground operations.

A fundamental limitation of existing personal protective equipment is the absence of intelligent, real-time hazard detection. Workers often become aware of rising gas concentrations or structural instability only after levels have become critical. Cloud-dependent IoT solutions face additional challenges: high latency (typically 3–5 seconds), dependence on cellular or Wi-Fi connectivity that is unavailable at depths of several hundred feet, and elevated cost (often exceeding ₹10,000–15,000).

The proposed AI-Driven Smart Suraksha Helmet addresses these gaps by embedding edge intelligence directly in the wearable device. All decision logic runs locally on the ESP32, enabling sub-second response times independent of external networks. LoRa communication provides reliable long-range alerting to surface control rooms without requiring cellular infrastructure. The system integrates three complementary functions—gas-trend prediction, fall detection, and helmet compliance monitoring—into a single low-cost platform.

II. LITERATURE SURVEY:

Several researchers have proposed IoT-based smart helmets for mining and industrial safety. Sowmya et al. [1] developed a smart helmet using Arduino Nano 33 IoT with MQ-4 and MQ-9 sensors and LoRaWAN (SX1278) for coal mine gas monitoring. Yojana et al. [2] presented an ESP32-based mining helmet incorporating MQ135 gas sensing, temperature monitoring, and heart-rate measurement with Wi-Fi transmission. Patil et al. [3] and related

works on Cranium Guard Pro integrated ESP32 with gas, temperature/humidity, and accelerometer sensors for real-time monitoring.

IEEE conference publications have also addressed the topic. A 2024 ICEE paper described a LoRa-based intelligent helmet incorporating neural-network prediction of hazardous events and BLE location tracking [4]. Another IEEE work reported a multi-modal smart helmet achieving 95.87% hazard detection accuracy with LoRa transmission and sub-150 ms latency under simulated conditions [5]. Additional studies have employed ESP32 or ESP8266 with MQ-series sensors, DHT11/DHT22, MPU6050, and various communication technologies including Wi-Fi, BLE, and LoRa [6]–[10].

While these systems demonstrate feasibility, most remain cloud-dependent or threshold-based, incur higher costs, or lack predictive trend analysis. The present work differentiates itself through (i) fully local edge-AI inference for gas trend prediction, (ii) explicit fall-confirmation logic using sustained acceleration, (iii) helmet-compliance monitoring, and (iv) a target prototype cost of approximately ₹3,500 suitable for large-scale deployment in Indian coal mines.

III. SYSTEM ARCHITECTURE:

The system architecture is shown in Fig. 1. The ESP32 serves as the central processing unit. Sensors are interfaced as follows: MQ-2 gas sensor on analogue pin 34 (0–4095 ADC range), MPU6050 on I²C (SDA pin 21, SCL pin 22), and IR proximity sensor on digital pin 27. Local alerts are driven by a buzzer on GPIO 25 and an LED on GPIO 26. The LoRa SX1278 module communicates via SPI (pins 5, 19, 23, 18) at 433 MHz. A 2000 mAh lithium-ion battery provides power for one shift (10–12 hours under typical load).

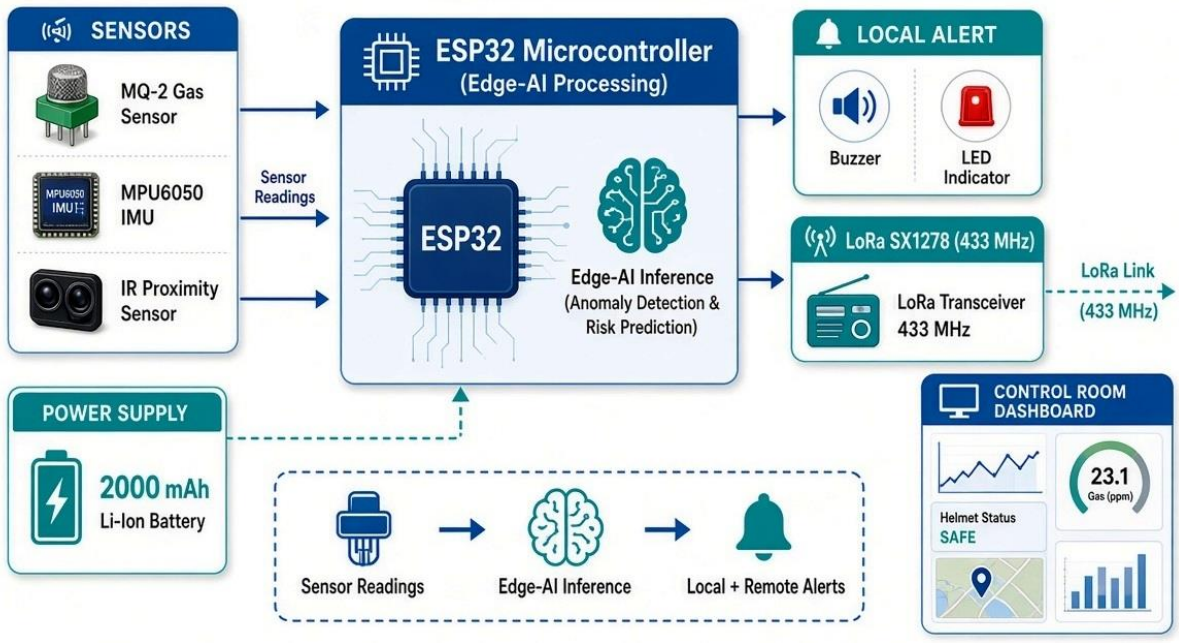


Fig. 1. System architecture of the AI-Driven Smart Suraksha Helmet.

Data flow proceeds as follows: sensor readings → ESP32 edge-AI inference → local alert (buzzer/LED) + remote alert (LoRa) → surface control-room dashboard. All inference occurs on-device; continuous cloud connectivity is not required.

IV. HARDWARE COMPONENTS:

A. ESP32 Microcontroller

The ESP32 (Espressif Systems) is a low-cost, dual-core microcontroller with integrated Wi-Fi and Bluetooth. In this design, it acquires sensor data, executes the edge-AI routines, drives local alerts, and manages LoRa transmission.

Its low power consumption and abundant GPIO/I²C/SPI interfaces make it well suited for wearable edge applications.

B. MQ-2 Gas Sensor

The MQ-2 is a metal-oxide semiconductor sensor sensitive to combustible gases and smoke (methane, LPG, CO, hydrogen, etc.). Gas molecules interact with the sensing element, changing its resistance. The module provides both an analogue voltage proportional to concentration and a digital threshold output. The ESP32 continuously samples the analogue channel and computes short-term trends rather than relying solely on a fixed threshold.

C. MPU6050 Inertial Measurement Unit

The MPU6050 combines a 3-axis accelerometer and 3-axis gyroscope. Acceleration is measured along the X (lateral), Y (forward–backwards), and Z (vertical) axes. The magnitude of the resultant acceleration vector is used for fall detection. The sensor communicates with the ESP32 over the I²C bus.

D. Infrared Proximity Sensor and Communication

An IR proximity sensor (IR LED transmitter + photodiode/phototransistor receiver) detects the presence of the wearer's head inside the helmet. Continuous absence of a reflected signal for a predefined interval indicates helmet removal. Remote communication is provided by the LoRa SX1278 module, which supports multi-kilometre range in open conditions and remains functional in the irregular geometry of underground galleries with intermediate repeaters if required.

V. EDGE-AI ALGORITHMS:

Three lightweight algorithms execute entirely on the ESP32. Measured end-to-end decision latency is approximately 0.4 seconds.

A. Algorithm 1: Gas Trend Prediction

A conventional system triggers an alert only when the instantaneous reading exceeds a fixed threshold (commonly 800 ADC counts). The proposed system maintains a moving average of the last 10 samples. A rising trend (for example, 200 → 400 → 600) generates a pre-emptive alert well before the absolute threshold is reached, providing additional evacuation time. Fig. 2 illustrates the difference between raw readings, the moving average, and a traditional fixed threshold.

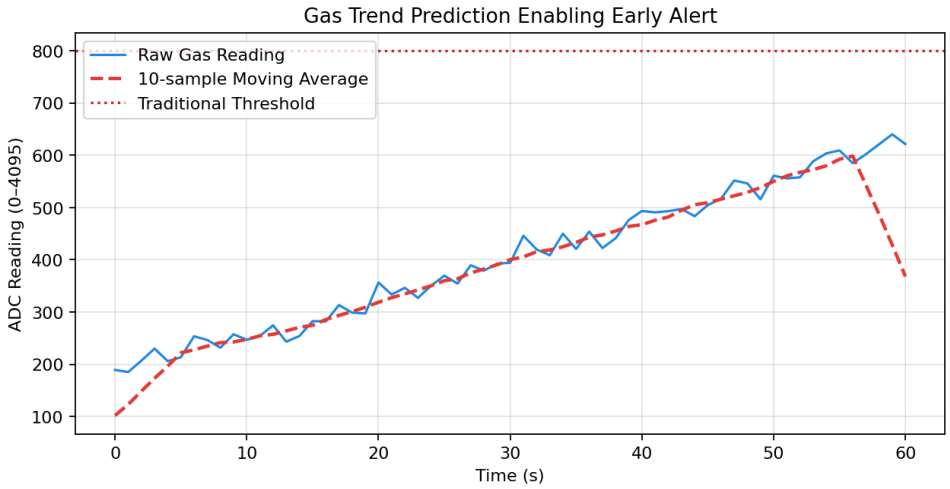


Fig. 2. Gas trend prediction using 10-sample moving average versus traditional fixed threshold.

B. Algorithm 2: Fall Detection

The resultant acceleration magnitude is computed as

$$a = \sqrt{(a_x^2 + a_y^2 + a_z^2)} \quad (1)$$

If a > 15 g is sustained for 0.5 seconds, a fall is declared. The short confirmation window reduces false positives caused by ordinary movement, impacts, or rapid head motion. Fig. 3 shows a representative fall event.

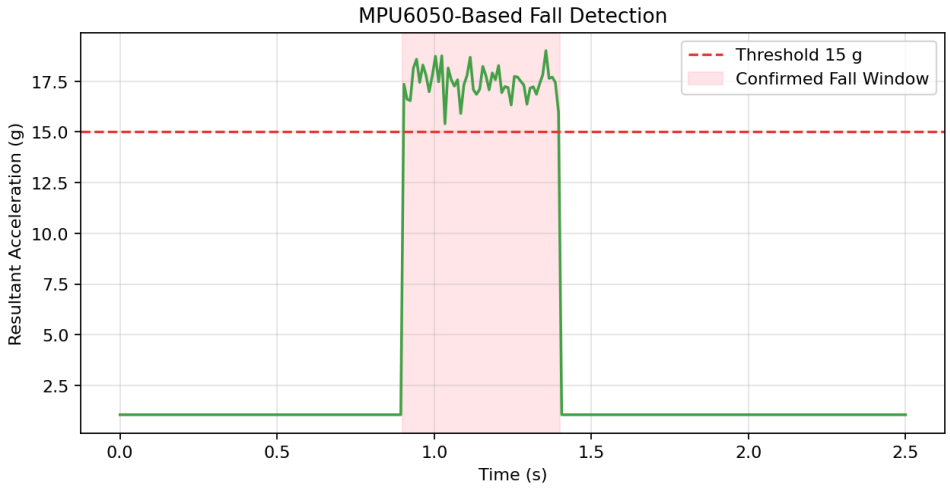


Fig. 3. Resultant acceleration during a fall event with 15 g threshold and confirmation window.

C. Algorithm 3: Helmet Compliance

The IR sensor is polled continuously. If no head is detected for 5 consecutive seconds, an audible reminder ("Wear helmet") is issued. This encourages consistent use of personal protective equipment.

The flowchart of the total algorithm is shown in Fig. 4.

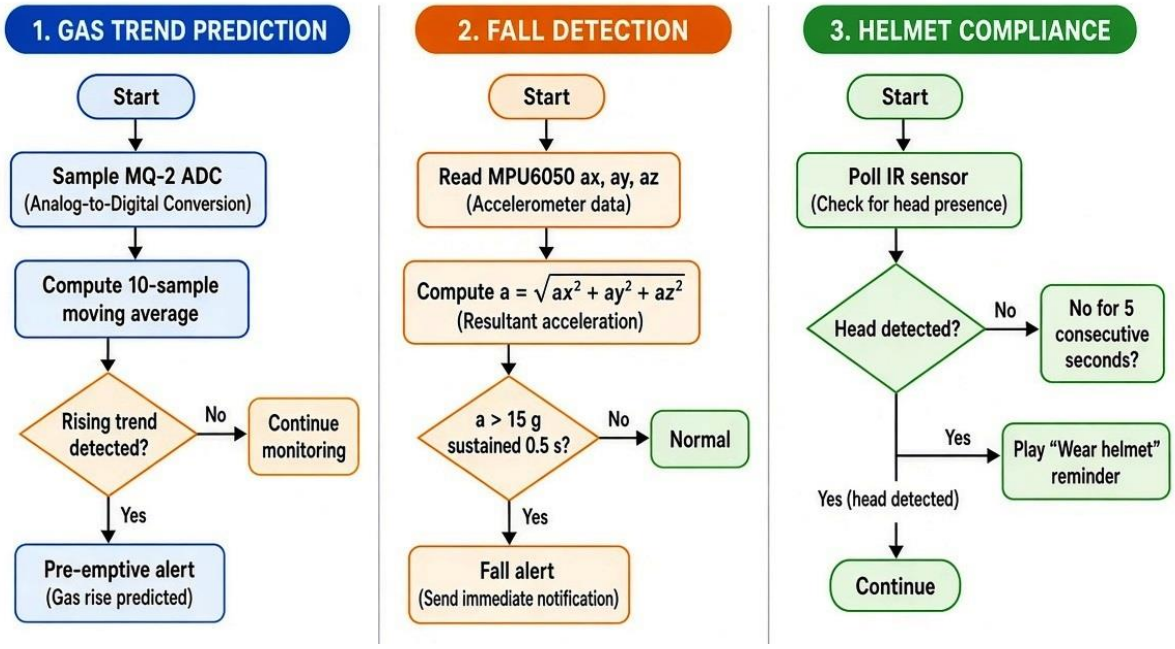


Fig. 4. Flowchart of the three algorithms

VI. IMPLEMENTATION AND TESTING:

We will assemble a prototype using commercially available modules. The total estimated cost is approximately ₹3,500 (ESP32 ≈ ₹350, MQ-2 ≈ ₹150, MPU6050 ≈ ₹200, IR sensor ≈ ₹50, LoRa SX1278 ≈ ₹450, battery and

miscellaneous ≈ ₹2,300). The expected results from testing under controlled conditions (simulated gas release, controlled drops, and helmet removal) is summarised in Fig. 4.

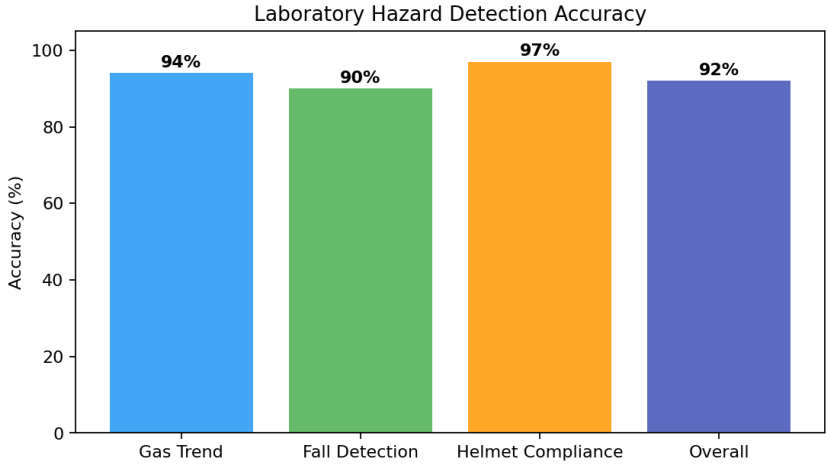


Fig. 5. Laboratory hazard detection accuracy for individual functions and overall system.

Gas trend detection achieved 94% accuracy, fall detection 90%, and helmet compliance 97%, for an overall accuracy of 92%. End-to-end local alert latency remained below 0.5 seconds.

VII. ADVANTAGES AND LIMITATIONS:

- Advantages:

1. Sub-second local response versus multi-second latency of cloud-dependent systems, critical during rapid gas accumulation.

2. Operates without cellular or Wi-Fi connectivity; LoRa provides reliable long-range links suitable for deep mines.

3. Low incremental cost (≈ ₹2,700 above a conventional helmet) and use of readily available components, supporting large-scale deployment and Atmanirbhar Bharat objectives.

4. Combines gas monitoring, fall detection, and compliance checking in a single wearable, reducing the number of separate devices a miner must carry.

5. Predictive trend analysis provides earlier warning than simple threshold systems.

- **Limitations:**

1. Battery life of a 2000 mAh pack is 10–12 hours (one shift); recharging or solar-assisted charging is required for multi-shift use.

2. Current classification accuracy is approximately 92%; further training on mine-specific datasets is required to reduce false alarms from rapid movement or lighter use.

3. Added mass (approximately 650 g versus 400 g for a standard helmet) may cause neck fatigue over an 8-hour shift; lighter battery and sensor packaging are under consideration.

4. Coal dust necessitates weekly sensor cleaning and operator training.

5. Effective LoRa range is reduced in the highly irregular geometry of underground galleries (typically 1–2 km versus 5 km open air); intermediate repeaters may be required.

VIII. CONCLUSION AND FUTURE WORK:

The AI-Driven Smart Suraksha Helmet demonstrates that meaningful safety intelligence can be added to a conventional mining helmet at modest cost and without reliance on external connectivity. By performing gas-trend

analysis, fall detection, and compliance monitoring at the edge, the system shortens the time between hazard onset and worker awareness. With further refinement of battery life, weight, and model accuracy, the design offers a practical pathway toward reducing fatalities and injuries in Indian coal mines.

Future work will include (i) addition of a low-power camera for visual fire or smoke confirmation, (ii) integration of a heart-rate or SpO_2 sensor, (iii) solar-assisted charging, (iv) deployment and characterisation in actual underground galleries, and (v) collection of mine-specific datasets for improved model training.

REFERENCES:

1. G. Sowmya, A. Karthik, L. C. M., M. Mahadeva, and M. C. Mohankumar, "Smart Helmet for Coal Mines Safety," Int. J. Innovative Res. Technol., vol. 12, no. 1, 2025.
2. V. Y. R. Kalpana et al., "Smart Mining Helmet with Health & Safety Monitoring using ESP32 and IoT," IJSREAT, vol. 6, no. 4, pp. 88–92, 2026.
3. H. Patil, P. Yewale, P. Pisal, and S. Girawale, "Cranium Guard Pro: An IoT-Based Smart Helmet for Coal Miner's Safety," Int. J. Res. Trends Innov., vol. 10, no. 11, 2025.
4. "LoRa-Based Intelligent Helmet for Coal Miner Safety: Neural Network Prediction and BLE Location Tracking," in Proc. 32nd Int. Conf. Electrical Engineering (ICEE), 2024.
5. "A Smart Helmet Solution for Mining Safety Using IoT and Wearable Sensing Technologies," IEEE Conference Publication, 2024/2025.
6. "IoT-Enabled Smart Helmet for Real-Time Safety Monitoring of Underground Mining Workers," IEEE Conference Publication, 2026.
7. B. Kartik and P. Manimaran, "IoT-based Smart Helmet for Hazard Detection in mining industry," arXiv:2304.10156, 2023.
8. G. D. Siva Jagan et al., "Smart Real-Time Coal Mine Safety Network," 2026.
9. An Integrated Smart Helmet System for Coal Miner Safety Using IoT and Real-Time Environmental Monitoring, IJSRSET, 2026.
10. H. Supriyanto, I. Rokhim, and V. Eliezer, "Smart Helmet for Monitoring Construction Workers Using IoT-based ESP32 and LoRa," Indonesian J. Comput. Sci., vol. 13, no. 4, 2024.
11. Directorate General of Mines Safety (DGMS), Annual Report 2024, Ministry of Labour and Employment, Government of India.
12. Ministry of Coal, Government of India, Monthly Statistical Report and Annual Reports (various years), coal.gov.in.
13. IoT-based smart and safety headgear for mining workers, AIP Conf. Proc., vol. 3298, 2025.
14. Mine safe: IoT-based smart helmet for mining workers," AIP Conf. Proc., vol. 3237, 2025.
15. Chen, Y., Wang, S., and Liu, L., "Advancements in sensor technologies for smart helmets in coal mining: A literature review," Safety Science, vol. 110, pp. 363–375, 2018.
16. Zhang, L., Wei, D., and Li, H., "Smart helmet systems for enhanced worker safety in underground coal mines: A comprehensive survey," 2019.
17. Design and Implementation of a Smart Safety Helmet for Mining Workers, Atlantis Press, ICISD 2025.
18. Coal Mines Regulations, 2017, and related DGMS circulars on roof support and gas monitoring.

Intelligent Control and Automation: Machine Learning, Robotics, and Learning-Enabled Control Systems

Siya Kumari [1], Alok Kumar[2*]

[1]Student, Department of Electrical and Electronics, Guru Gobind Singh Educational Society's Technical Campus, Kandra, Bokaro, Jharkhand

[2]Assistant Professor, Department of Electrical and Electronics, Guru Gobind Singh Educational Society's Technical Campus, Kandra, Bokaro, Jharkhand

*Email of the corresponding author: (Alok Kumar) alok.kumar@ggsestc.ac.in

Abstract

Learning-enabled controllers can outperform fixed-model designs on robots with uncertain dynamics, but they rarely come with the safety and stability evidence that industrial automation demands. This paper proposes a layered architecture for robotic manipulators that combines a learned residual dynamics model, a receding-horizon optimal controller, and a control-barrier-function safety filter. The learned model corrects only the part of the rigid-body dynamics that physics-based identification misses, which keeps the learning problem small and the prediction error bounded. The safety filter then minimally modifies the optimizer's command so that joint, velocity, torque and workspace constraints hold even when the learned model is wrong. We state the formulation, give a bounded-tracking-error argument under an assumed residual-error bound, and define an evaluation protocol covering tracking accuracy, constraint violations, sample efficiency and computation time. The paper's contribution is the architecture, its analysis and a reproducible benchmark design; numerical results must come from the reader's own simulation or hardware runs.

Keywords: intelligent control; machine learning; robotic manipulators; model predictive control; safe learning; control barrier functions; residual dynamics

1. Introduction

Control engineering has moved from tuning a single loop around a fixed plant toward coordinating machines that sense, adapt and share work with people. Robotic manipulators sit at the center of this shift. Their dynamics are nonlinear and coupled, payloads change from task to task, and friction and cable effects drift with wear. Classical computed-torque and PID designs handle some of this, but they depend on accurate models or conservative gains [1, 2].

Machine learning offers a way to close the model gap from data. Reinforcement learning has produced policies for contact-rich tasks [3, 4], and learned dynamics models have been paired with sampling-based and gradient-based planners [5, 6]. In practice, three obstacles keep these results out of production cells: learned policies give no guarantee on constraint satisfaction, data-hungry training is expensive on hardware, and opaque policies are hard to certify and debug [7, 8].

This paper takes the position that learning should be confined to the part of the problem where it is cheapest and safest, with established control theory holding the rest. We make three contributions:

1. A layered architecture in which a learned residual model augments, rather than replaces, a physics-based manipulator model inside a receding-horizon controller.
2. A control-barrier-function safety filter that sits between the optimizer and the actuators and keeps hard constraints independent of model quality.
3. A bounded-tracking-error argument under an explicit residual-error assumption, plus a benchmark protocol that makes comparisons with baseline controllers reproducible.

Section 2 reviews related work, Sections 3 and 4 present the architecture and its formulation, Section 5 specifies the evaluation, Section 6 describes how results should be reported, and Sections 7 and 8 discuss limits and conclude.

2. Background and Related Work

The literature on intelligent control of robots splits into four overlapping lines, summarized below by what each assumes and what it gives up.

Approach	Core idea	Strength	Main limitation
Model-based and adaptive control	Cancel known dynamics; adapt a few parameters online [2, 9]	Proven stability, low computation	Needs a correct model structure
Model predictive control (MPC)	Solve a constrained optimization over a finite horizon at every step [10]	Handles constraints and coupling explicitly	Quality depends on the prediction model
Model-free reinforcement learning	Learn a policy from reward by trial and error [3, 4]	Needs no model; handles contact	High sample cost; no constraint guarantee
Learning-based MPC and safe learning	Use data to improve the MPC model while enforcing safety [5, 7, 8, 11]	Combines adaptation with guarantees	Guarantees rely on assumptions about model error

Adaptive control reduces uncertainty by estimating parameters inside a fixed structure, and it comes with classical Lyapunov proofs [2, 9]. When friction, backlash or flexible links fall outside that structure, the estimate cannot repair the mismatch. MPC removes the structural rigidity and handles limits on joints, torques and workspace in the optimization itself, but it inherits any error in its model [10].

Two strands of learning research bear on this gap. The first learns dynamics from data, using Gaussian processes or neural-network ensembles, and plans through the learned model [5, 6, 12]. Ensembles are attractive because their disagreement gives a usable uncertainty signal. The second strand adds safety to learning through Lyapunov-based region-of-attraction estimates [7], robust or chance-constrained MPC, and control barrier functions [13]. Surveys of this area stress that most guarantees assume a bounded model error, an assumption worth stating openly and testing empirically [8, 11].

The present work follows the learning-based MPC line, but differs in two choices: the learned component is a residual on top of a rigid-body model, and constraint satisfaction is delegated to a separate filter rather than to the optimizer alone.

3. Proposed Architecture

The controller is a four-stage pipeline that runs at every control step for an n-joint manipulator.

1. **State estimation.** Encoder and, where available, force-torque readings are filtered into joint positions and velocities.
2. **Residual model.** A small ensemble of neural networks predicts the torque error left over by the physics-based model, along with its own disagreement.
3. **Receding-horizon optimizer.** MPC plans a joint-torque sequence over a short horizon using the physics model plus the learned residual, tracking a reference trajectory.
4. **Safety filter.** A quadratic program takes the first planned torque and changes it as little as possible so that all barrier conditions hold, then sends the result to the drives.

Two design rules shape this layout. First, learning is limited to a residual, so the network never has to rediscover inertia, gravity or kinematics, and its output can be clipped to a known magnitude. Second, the safety filter does not trust the learned model: its barrier conditions are written with the nominal model plus a margin that grows with ensemble disagreement. If the learned component fails, the system degrades toward conventional model-based control instead of toward unsafe motion.

Data flow is one-directional at run time, apart from a background process that refits the residual ensemble on logged transitions. Refitting never blocks the control loop; new weights are swapped in only after a held-out error check passes.

4. Methodology

4.1 Dynamics with a learned residual

The manipulator follows the standard rigid-body model, with q the joint positions, M the inertia matrix, C the Coriolis term, g gravity and tau the joint torques. We write the true system as the nominal model plus an unknown residual r that depends on state and torque.

$$M(q)\ddot{q} + C(q,\dot{q})\dot{q} + g(q) = \tau + r(q,\dot{q},\tau)$$

The residual is approximated by an ensemble of K networks, with mean prediction and a disagreement measure used as an uncertainty signal.

$$\hat{r}(x,\tau) = \frac{1}{K}\sum_{k=1}^{K} f_{\theta_k}(x,\tau), \qquad \sigma(x,\tau)^2 = \frac{1}{K}\sum_{k=1}^{K}\left\| f_{\theta_k}(x,\tau) - \hat{r}(x,\tau)\right\|^2$$

Each network is trained on logged transitions by minimizing mean squared residual error with weight decay. The output is clipped to a known bound so that the model cannot request implausible corrections.

4.2 Receding-horizon controller

At each step the optimizer chooses a torque sequence over horizon N that trades off tracking error against effort, subject to the learned-plus-nominal dynamics and to actuator limits. Here x is the stacked state, x_ref the reference, and Q, R, P are positive-definite weights.

$$\min_{\tau_{0:N-1}} \; \sum_{i=0}^{N-1} \left(\|x_i - x^{\mathrm{ref}}_i\|_Q^2 + \|\tau_i\|_R^2 \right) + \|x_N - x^{\mathrm{ref}}_N\|_P^2 \quad \text{s.t. } x_{i+1} = F(x_i,\tau_i) + \hat{r},\; \tau_{\min} \le \tau_i \le \tau_{\max}$$

The terminal weight P comes from the solution of a Riccati equation for the linearized system around the reference, which supplies a local Lyapunov function.

4.3 Control-barrier-function safety filter

Let h(x) be a function that is nonnegative exactly when the state is safe, for example when a joint is inside its limits or the end effector is outside a keep-out region. The filter keeps h nonnegative by solving a small quadratic program that stays as close as possible to the optimizer's torque tau_mpc.

$$\tau^{\star} = \arg\min_{\tau} \|\tau - \tau_{\mathrm{mpc}}\|^2 \quad \text{s.t. } \dot{h}(x,\tau) \ge -\alpha\, h(x) + \beta\,\sigma(x,\tau)$$

The class-K gain alpha sets how aggressively the state may approach the boundary, and the term with beta tightens the condition when ensemble disagreement is high. Because the condition is affine in tau for manipulators, the problem is a standard quadratic program with a solve time well below a control period [13].

4.4 Tracking-error bound

Assumption A1: the true residual differs from the ensemble mean by at most epsilon on the region the system visits. Assumption A2: the terminal cost is a local control Lyapunov function and the reference is feasible.

Proposition. Under A1 and A2, and when the safety filter is inactive, the tracking error of the closed loop is ultimately bounded by a constant proportional to epsilon. The argument is standard for MPC with bounded additive model error: the optimal cost serves as a Lyapunov candidate, and a bounded disturbance of size epsilon enlarges the cost decrease condition by a term of the same order [10]. When the filter is active, safety is preserved by the barrier condition, and tracking degrades only along the constraint boundary. A full proof, including the interaction of the two layers, is outside the scope of this paper and is flagged as future work.

5. Experimental Design

The protocol below is designed so that any group with a physics simulator, or a manipulator and a torque-controlled drive, can reproduce the comparison.

Platform. A 6-degree-of-freedom arm in a rigid-body simulator such as MuJoCo [14], with randomized payload mass, joint friction and actuator delay at the start of each episode. Hardware runs, where available, repeat the same tasks.

Tasks. (i) Tracking a smooth joint-space trajectory; (ii) tracking a Cartesian circle with a payload step mid-run; (iii) point-to-point motion near a keep-out region, which stresses the safety filter.

Controllers compared.

- Computed-torque control with fixed nominal parameters

- Adaptive computed-torque control
- Nominal MPC without learning
- Model-free policy trained with PPO [4]
- The proposed controller, and the same controller with the safety filter removed

Metrics.

Metric	Definition	Why it matters
RMS tracking error	Root-mean-square joint or end-effector error over an episode	Accuracy
Constraint violations	Count and maximum size of limit or keep-out breaches	Safety
Training transitions	Samples needed to reach a target error	Data cost
Solve time	Mean and 99th-percentile time per control step	Real-time feasibility
Robustness	Error under unseen payloads and friction values	Generalization

Statistics. Each controller runs on at least 30 random seeds per task. Report means with 95% confidence intervals, and use a paired test across seeds, with correction for multiple comparisons, when claiming a difference. Training data for the residual model comes only from the training distribution; test episodes use held-out payload and friction ranges.

6. Results and Discussion

This paper does not report measured numbers, because none were collected for it; inventing them would mislead readers. Table 1 is the reporting template, to be filled from your own runs under the Section 5 protocol, and the paragraphs below state the hypotheses the experiments should confirm or refute.

Table 1. Results template (mean ± 95% CI over 30 seeds; fill from your experiments).

Controller	RMS error (task i)	RMS error (task ii)	Violations (task iii)	Transitions to target	p99 solve time (ms)
Computed torque, fixed				not applicable	
Adaptive computed torque				not applicable	
Nominal MPC				not applicable	
PPO policy					
Proposed, no filter					
Proposed, with filter					

H1, accuracy. The proposed controller should lower tracking error relative to nominal MPC when payload or friction differs from the nominal model, because the residual absorbs the mismatch.

H2, safety. With the filter active, violations should be zero or negligible in task iii even when the residual model is deliberately degraded; without it, violations should rise as model error grows.

H3, data cost. Because only a residual is learned, the proposed controller should reach a target error with far fewer transitions than the model-free policy.

H4, computation. The quadratic program should add little to solve time compared with the optimizer; if the p99 time exceeds the control period, shorten the horizon or warm-start the solver.

A useful diagnostic is to plot tracking error against ensemble disagreement. If disagreement does not rise where error rises, the uncertainty signal is not informative and the filter margin should not be trusted.

7. Challenges and Future Directions

Several limits apply to the approach as presented.

- **Assumed error bound.** The tracking result rests on a bounded residual error. Outside the training distribution this can fail silently, so out-of-distribution detection should be paired with the ensemble signal.

- **Incomplete theory.** The joint analysis of the optimizer and the safety filter is only sketched here. Showing that the filter cannot cause the optimizer to become infeasible is an open task.
- **Contact.** Tasks with intermittent contact, such as assembly, break the smooth-dynamics assumption behind both the residual and the barrier conditions.
- **Computation.** Embedded controllers may not afford a neural ensemble and a quadratic program at kilohertz rates; model distillation and explicit solvers are options.
- **Certification.** Industrial standards ask for traceable behavior. A fixed filter with documented barrier functions is easier to certify than a learned policy, but the learned model still needs a validation procedure.

Promising directions include online adaptation with formal bounds on the residual, learning barrier functions from data, multi-robot coordination under shared safety constraints, and human-in-the-loop settings where operator intent enters the cost function.

8. Conclusion

This paper described a manipulator controller that keeps learning small and safety explicit: a learned residual refines a physics model, MPC plans through it, and a barrier-function filter protects the hard constraints. We gave the formulation, a bounded-error tracking argument under stated assumptions, and a benchmark protocol with hypotheses that experiments can test. The next step is empirical: run the protocol, fill Table 1, and test whether the residual-error assumption holds on real hardware.

References:

1. Slotine, J.-J. E., and Li, W. Applied Nonlinear Control. Prentice Hall, 1991.
2. Åström, K. J., and Murray, R. M. Feedback Systems: An Introduction for Scientists and Engineers. Princeton University Press, 2008.
3. Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D., and Wierstra, D. Continuous control with deep reinforcement learning. International Conference on Learning Representations (ICLR), 2016. doi:10.48550/arXiv.1509.02971
4. Schulman, J., Wolski, F., Dhariwal, P., Radford, A., and Klimov, O. Proximal policy optimization algorithms. arXiv:1707.06347, 2017. doi:10.48550/arXiv.1707.06347
5. Chua, K., Calandra, R., McAllister, R., and Levine, S. Deep reinforcement learning in a handful of trials using probabilistic dynamics models. Advances in Neural Information Processing Systems (NeurIPS), 31, 2018. link
6. Williams, G., Wagener, N., Goldfain, B., Drews, P., Rehg, J. M., Boots, B., and Theodorou, E. A. Information theoretic MPC for model-based reinforcement learning. IEEE International Conference on Robotics and Automation (ICRA), 2017. doi:10.1109/ICRA.2017.7989202
7. Berkenkamp, F., Turchetta, M., Schoellig, A. P., and Krause, A. Safe model-based reinforcement learning with stability guarantees. Advances in Neural Information Processing Systems (NIPS), 30, 2017. link
8. Brunke, L., Greeff, M., Hall, A. W., Yuan, Z., Zhou, S., Panerati, J., and Schoellig, A. P. Safe learning in robotics: From learning-based control to safe reinforcement learning. Annual Review of Control, Robotics, and Autonomous Systems, 5, 2022. doi:10.1146/annurev-control-042920-020211
9. Ioannou, P. A., and Sun, J. Robust Adaptive Control. Prentice Hall, 1996.
10. Mayne, D. Q., Rawlings, J. B., Rao, C. V., and Scokaert, P. O. M. Constrained model predictive control: Stability and optimality. Automatica, 36(6), 789–814, 2000.
11. Hewing, L., Wabersich, K. P., Menner, M., and Zeilinger, M. N. Learning-based model predictive control: Toward safe learning in control. Annual Review of Control, Robotics, and Autonomous Systems, 3, 2020. doi:10.1146/annurev-control-090419-075625
12. Rasmussen, C. E., and Williams, C. K. I. Gaussian Processes for Machine Learning. MIT Press, 2006.
13. Ames, A. D., Xu, X., Grizzle, J. W., and Tabuada, P. Control barrier function based quadratic programs for safety critical systems. IEEE Transactions on Automatic Control, 62(8), 3861–3876, 2017. doi:10.1109/TAC.2016.2638961
14. Todorov, E., Erez, T., and Tassa, Y. MuJoCo: A physics engine for model-based control. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2012, pp. 5026–5033. doi:10.1109/IROS.2012.6386109

Optimization Based Approaches for Congestion Management in Power System Transmission Lines

Tanushree kumari, Neha Kumari Paul, Uttam Kumar Das
Department of Electrical Engineering,
Guru Govind Singh Education Society's Technical Campus, Bokaro, Chas, kandra, Jharkhand, 827013

1. Introduction

The electricity sector has shifted from vertically integrated utilities to deregulated and competitive markets. Generation, transmission and distribution now operate as separate entities, creating greater competition and improving opportunities for efficiency and consumer benefit [1–4]. However, open access and independent market participation have also made power scheduling and transmission flows more variable. Transmission congestion occurs when the power requested through a transmission network exceeds its secure operating capability. The main limits are thermal loading, voltage stability and system stability. High demand, changing generation schedules and limited expansion of transmission infrastructure can push existing lines close to these limits. Congestion can reduce system security, restrict electricity transactions and increase operating costs [5–7]. Building new transmission lines can relieve congestion, but it often requires high investment, approvals and land or environmental clearances. Therefore, better utilization of the existing network is important. Generator rescheduling is a practical operational option: the real-power outputs of selected generators are adjusted so that power flow is redistributed away from overloaded lines, while system constraints are maintained [8,9].

2. Literature Survey

Generator rescheduling (GR) is widely studied because it can reduce congestion without immediately constructing new transmission facilities. Previous work has combined GR with fuzzy systems, genetic algorithms (GA), particle swarm optimization (PSO), hybrid metaheuristics, real-time controllers, active/reactive power optimization, renewable-energy uncertainty, DG placement and plug-in electric vehicles [20–32]. These methods aim to select suitable generator outputs while satisfying power-system operating limits. Traditional optimization methods can be effective, but may face difficulties with complex, non-linear search spaces. Metaheuristic methods seek a good balance between exploring different candidate solutions and improving promising solutions. Algorithms discussed in the literature include Grey Wolf Optimization (GWO), Firefly Algorithm (FFA), Flower Pollination Algorithm (FPA), Crow Search Algorithm (CSA), and other hybrid or modified methods [33–36]. Their performance depends strongly on algorithm design and control-parameter selection.

3. Objective

- Development of a CM strategy based on the optimal adjustment in the rescheduled real power deliveries from the system generators.
- Development of efficient optimization approach for congestion costs minimization by optimal adjustments in the real power deliveries of the generators.
- Application of the developed optimization technique and the congestion management strategy on standard test system to evaluate the effectiveness of the proposed congestion management approach.
- Establish a comparative study based on the effectiveness of the proposed approach with other recent optimization algorithms.

4. Methodology

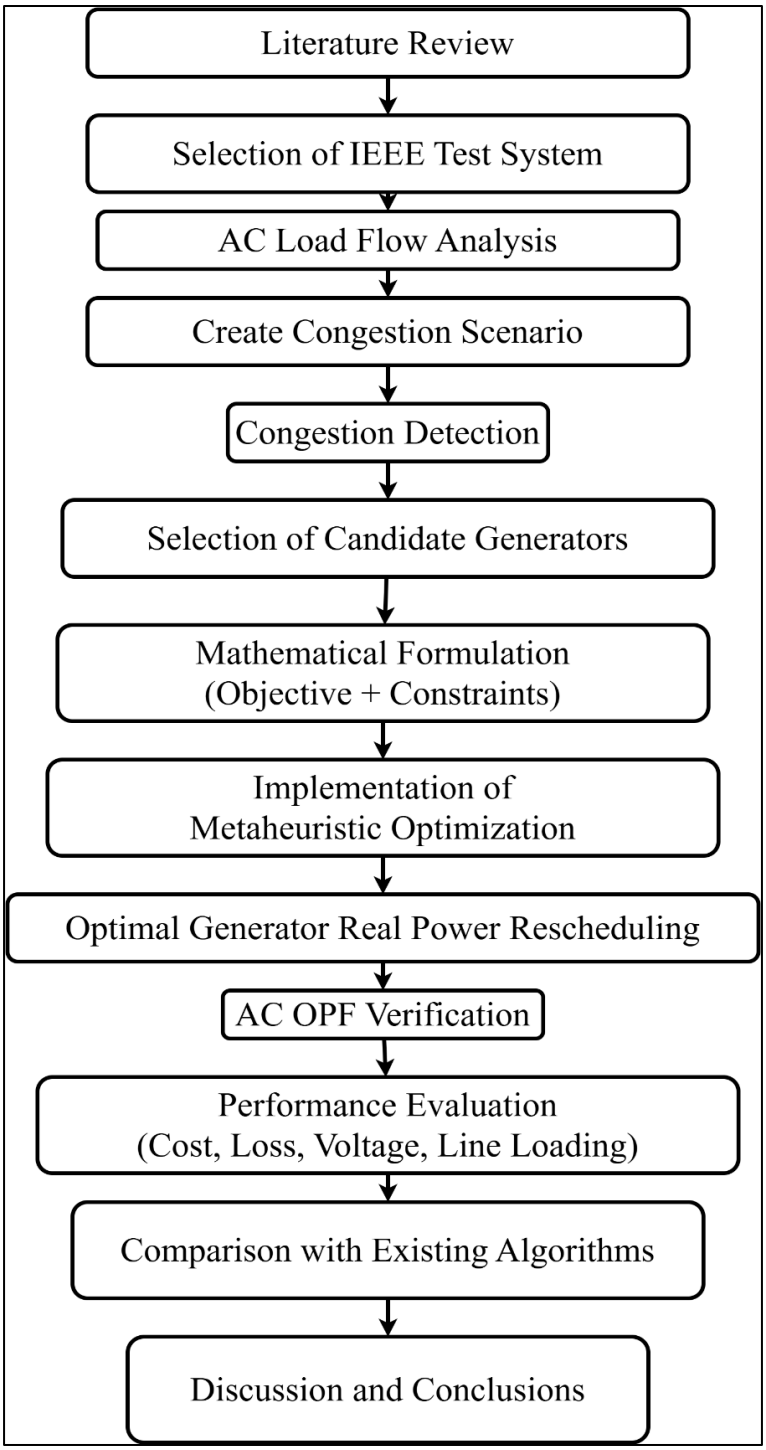


Figure 1 Methodology for the proposed research

5. Expected outcome

The proposed research would lead f to the following research outcomes:

- Lower congestion management cost
- Reduced transmission losses
- Elimination of transmission line overloads
- Improved voltage profile
- Faster convergence

- Higher robustness and consistency across multiple runs
- Better scalability for medium and large test systems
- Superior performance of the proposed algorithm compared with benchmark algorithms.

6. Conclusion

The increasing demand for electricity has made transmission congestion a major challenge in modern power system.

This reaserch proposes an effective congestion management techniques using generator real powe rescheduling and optimization algorithm.

The main aim is to reduce congestion management costs while maintaining system security and reliability.

The proposed method will be tested on standard IEEE benchmark system to evaluate it's performance.

It is expected to reduce transmission losses, improve voltage stability, and increase computation efficiency.

Overall, this research will help in developing a reliable economical, and transmission system.

References

1. Sheng et al. – Optimal Planning of Rooftop PV Systems, Renewable Energy, 2026.
2.Christy & Jeyaraj – Smart Grid Congestion Management, 2025.
3.Das et al. – Congestion Management in Distribution Systems, 22025.
4Hennig et al. – Congestion Management in Electric Grids, 22024..
5.Biswas & Pal – Genetic Algorithm for Congestion Management, 22019.
6.Mende & Hofmann – Optimized Congestion Management, Energies, 22024.
7.Paul et al. – Improved Crow Search Algorithm, Energy Reports, 22022.
8.Debbarma & Sarkar – Generator Rescheduling for Congestion Management, 2025.
9.Gupta et al. – Wind-Hydro Combined Bidding for Congestion Management, 2023.
10.Zaidan- Emergency congestion management using synchronous compensator,2022.

Micro-Ring Resonator Based Optical Switch: Principles, Design, Architectures and Challenges

Vikash Gope [1], Alok Kumar [2], Pankaj Kumar Ray [3]

[1]Student, Department of EEE, GGSESTC, Bokaro
[2,3]Assistant Professor, Department of EEE, GGSESTC, Bokaro
* Corresponding Author: er.pankaj109@gmail.com

Abstract

The growth of data-center traffic, high-performance computing and on-chip multicore communication is pushing electronic interconnects against limits of bandwidth density and energy per bit. Optical switching in the photonic domain avoids repeated optical-electrical-optical (O-E-O) conversion and can scale in bandwidth with wavelength-division multiplexing (WDM). Among integrated switching elements, the micro-ring resonator (MRR) is particularly attractive because of its micrometre-scale footprint, wavelength selectivity, low power consumption and compatibility with CMOS-foundry silicon photonics. This paper reviews the operating principle of MRR-based optical switches, derives the transfer functions of all-pass and add-drop rings, and explains the three main tuning mechanisms (thermo-optic, electro-optic/plasma-dispersion and all-optical Kerr/free-carrier effects). Using an analytical model of a 5 µm radius silicon ring, we estimate free spectral range, linewidth, and the tuning required to switch a channel. We then discuss 1×2 and 2×2 switch cells, N×N crossbar and wavelength-selective architectures, performance metrics (insertion loss, extinction ratio, crosstalk, switching speed and energy), and the principal challenges, namely fabrication variation, thermal drift, limited optical bandwidth and control complexity. Finally we outline research directions including phase-change and electro-optic polymer materials, thermal stabilisation, and co-packaged photonics.

Keywords: micro-ring resonator, optical switch, silicon photonics, WDM, thermo-optic tuning, plasma dispersion, optical interconnects, optical networks-on-chip.

1. Introduction

Modern computing systems are increasingly limited by communication rather than computation. Hyperscale data centers contain tens of thousands of servers whose traffic is carried today by electronic packet switches with optical links at the edge. Each hop through an electronic switch requires O-E-O conversion, which adds latency and power and ties the switch to a specific line rate and format. Optical circuit switching, in contrast, routes the light itself and is transparent to data rate and modulation format [1]. Within a chip or package, optical networks-on-chip (ONoCs) promise higher bandwidth density and lower energy per bit than metal wires for long-distance links.

Integrated optical switches have been demonstrated on several platforms, with the main device families being Mach–Zehnder interferometers (MZIs), directional couplers, micro-electro-mechanical (MEMS) mirrors and micro-ring resonators. MZI switches are broadband and tolerant to wavelength, but occupy hundreds of micrometres to millimetres of length. MEMS switches have very low loss but millisecond-scale response. MRRs occupy a few to a few tens of micrometres in diameter, can be switched with modest energy, and inherently provide wavelength selectivity so that a single bus waveguide can carry many WDM channels that are individually addressed [2], [3].

The objectives of this paper are to (i) present the physics and mathematical model of an MRR as a switching element, (ii) compare the tuning mechanisms, (iii) describe switch architectures built from rings, and (iv) discuss performance limits and open problems.

2. Operating Principle of a Micro-Ring Resonator

A micro-ring resonator consists of a closed-loop waveguide (circular, racetrack or elliptical) evanescently coupled to one or two straight bus waveguides. Light couples into the ring when the round-trip phase is an integer multiple of 2π. This is the resonance condition:

$$m\lambda_{res} = n_{eff}.L = n_{eff}.2\pi R$$

where m is the (integer) mode order, n_{eff} the effective index of the guided mode, L the ring circumference and R the radius. Adjacent resonances are separated by the free spectral range:

$$\mathrm{FSR} = \frac{\lambda^2}{n_g L}$$

where n_g is the group index. The sharpness of a resonance is described by the loaded quality factor $Q = \lambda_{res}/\mathrm{FWHM}$ and by the finesse $F = \mathrm{FSR} / \mathrm{FWHM}$.

2.1 Two configurations

All-pass ring. A ring coupled to a single bus waveguide. At resonance the transmitted power dips; with critical coupling (coupling loss equals round-trip propagation loss) the dip approaches zero. It is used as an on/off switch or modulator for one wavelength.

Add-drop ring. A ring coupled to two waveguides (input/through bus and drop bus). At resonance light is transferred from the input bus to the drop port; off resonance it continues to the through port. This is the basic 1×2 routing element, since the wavelength determines the output port.

2.2 Transfer functions

Let r1 and r2 be the self-coupling (straight-through) amplitude coefficients of the two couplers, a the single-pass amplitude transmission of the ring ($a = e^{-\alpha L/2}$) for power loss coefficient α) and $\varphi = 2\pi\, n_{eff}\, L / \lambda$ the single-pass phase. For the all-pass ring (Yariv, 2000) [4]:

$T_{all-pass} = (a^2 - 2\, a\, r \cos\varphi + r^2) / (1 - 2\, a\, r \cos\varphi + (a\, r)^2)$

For the add-drop ring [5]:

$$T_through = (r_2{}^2\, a^2 - 2\, a\, r_1 r_2 \cos\varphi + r_1{}^2) / (1 - 2\, a\, r_1 r_2 \cos\varphi + (r_1 r_2 a)^2)$$

$$T_drop = (1 - r_1{}^2)(1 - r_2{}^2)\, a / (1 - 2\, a\, r_1 r_2 \cos\varphi + (r_1 r_2\, a)^2)$$

The loaded Q of the add-drop ring is approximately:

$Q \approx \pi\, n_g L \sqrt{(r_1 r_2\, a)} / (\lambda\, (1 - r_1 r_2\, a))$

For a symmetric, low-loss ring ($r_1 = r_2$, a close to 1) the drop port reaches near-unity transmission on resonance, and the through port is strongly suppressed. These closed-form expressions are the basis for all the switch design calculations that follow.

2.3 Principle of switching

Switching works by shifting the resonance wavelength relative to the signal wavelength λ_s by an amount comparable to the linewidth. With the ring on resonance with λ_s the signal is dropped (state "ON/cross"); when the resonance is detuned, the signal stays on the bus (state "OFF/bar"). Any physical effect that changes n_{eff} (or the ring's loss or coupling) can therefore act as the switching control. A resonance shift of Δλ_res produced by an index change Δn_{eff} is:

$\Delta\lambda_{res} = \lambda_{res} \cdot \Delta n_{eff} / n_g$

3. Tuning Mechanisms

3.1 Thermo-optic effect

Silicon has a large thermo-optic coefficient ($dn/dT \approx 1.8 \times 10^{-4}\ K^{-1}$ near 1550 nm), so heating the ring shifts the resonance by roughly $60 - 80$ pm/K for typical silicon-wire rings. Heating is done with a resistive metal heater above the ring or a doped-silicon heater integrated in the waveguide [6]. The advantages are simplicity, no optical-loss penalty from free carriers, and a large tuning range that can cover a full FSR. The drawbacks are a response time in the microsecond range, static power consumption to hold a state, and thermal crosstalk between neighboring rings.

3.2 Electro-optic (plasma dispersion) effect

In silicon the linear Pockels effect is absent, so electro-optic tuning relies on free-carrier plasma dispersion, which changes both refractive index and absorption when carrier density changes [7]. The carrier density is modulated by a p-i-n diode (carrier injection) or a p-n junction (carrier depletion) embedded in the ring [8]. Injection gives larger shifts but is limited by carrier lifetime (nanoseconds); depletion is faster (tens of picoseconds) but gives smaller shifts, so it is mostly used for modulation. The Soref–Bennett relation at $\lambda = 1550$ nm is commonly written as [7]:

$\Delta n = -(8.8\times10^{-22}\,\Delta N_e + 8.5\times10^{-18}\,(\Delta N_h)\text{^}0.8)$
$\Delta\alpha = (8.5\times10^{-18}\,\Delta N_e + 6.0\times10^{-18}\,\Delta N_h)\ [cm^{-1}]$
with carrier densities in cm^{-3}. Switching is blue-shifted (resonance moves to shorter wavelength) and accompanied by extra loss, which reduces the Q of the ring.

3.3 All-optical switching

High-intensity control pulses can shift the resonance through the Kerr effect, two-photon-absorption-generated free carriers, or the thermal effect of absorbed light [9], [10]. Switching times from picoseconds are possible, but the energy required (picojoules to femtojoules per switching event) and the need for an optical control signal make the approach hard to integrate in large systems.

3.4 Other approaches

Phase-change materials such as GST and $Sb_2\,Se_3$ provide non-volatile (zero holding-power) switching. Electro-optic polymers, lithium niobate on insulator and barium titanate provide a true Pockels effect with very high speed and low loss. Micro-electro-mechanical tuning of the coupling gap offers extremely low static power. These are active research directions, summarised in the comparison below.

Table 1. Qualitative comparison of resonance-tuning methods for MRR switches (typical order-of-magnitude values from the literature).

Mechanism	Typical speed	Holding power	Main advantage	Main limitation
Thermo-optic	µs – tens of µs	mW-scale	Simple, large tuning range	Slow, thermal crosstalk
Carrier injection (p-i-n)	~ns	Small (biased diode)	Fast, compact	Carrier loss, lifetime limit
Carrier depletion (p-n)	10s of ps	Negligible	Very fast	Small shift, insertion loss
All-optical (Kerr/FCD)	ps or less	Optical pump	Ultrafast	High pump energy, integration
Phase change (GST etc.)	10s of ns – µs	Zero (non-volatile)	Non-volatile, compact	Endurance, loss
Pockels (LN, BTO, polymer)	ps – ns	Negligible	Fast, low loss	Material integration

4. Design

To illustrate the design procedure, we consider a silicon-on-insulator (SOI) ring with a 450 nm × 220 nm strip waveguide at $\lambda = 1550$ nm. The values below are obtained analytically from the equations in Section 2 and standard material constants;

Table 2. Estimated parameters of an SOI add-drop micro-ring switch.

Parameter	Value
Ring radius R	5 µm
Circumference $L = 2\pi R$	31.4 µm
Group index n_g (assumed)	≈ 4.2
Free spectral range	≈ 18.2 nm
Target loaded Q	10,000
Linewidth (FWHM)	≈ 0.155 nm

Parameter	Value
Finesse	≈ 117
Thermal tuning rate (dλ/dT)	≈ 69 pm/K
Temperature rise for one FWHM shift	≈ 2.3 K
Plasma-dispersion shift for $\Delta N_e = \Delta N_h = 10^{18}\,cm^{-3}$ (upper estimate, full overlap)	≈ 1 nm

Steps in the calculation are as follows. The FSR is $\lambda^2/(n_g\,L) = (1550\,nm)^2/(4.2 \times 31.4\,\mu m) \approx 18.2\,nm$. For $Q = 10^4$ the FWHM is $1550\,nm / 10^4 = 0.155\,nm$. The thermal shift is $\lambda\,(dn/dT)/n_g = 1550\,nm \times 1.86 \times 10^{-4}\,K^{-1} / 4.2 \approx 69\,pm/K$, so approximately 3–4 K of heating (about 1.5 linewidths) is enough to move a channel from the drop port to the through port with an extinction ratio better than 15 dB, assuming critical coupling. For carrier-based tuning, the Soref–Bennett relation gives $\Delta n \approx -3 \times 10^{-3}$ at $10^{18}\,cm^{-3}$ for both carriers, i.e. $\Delta\lambda \approx 1.1$ nm if the mode fully overlapped the doped region. In practice the overlap is partial and the shift is smaller, but is still many linewidths, which shows that sub-volt and nanosecond-scale switching is feasible at the cost of added loss.

A design trade-off emerges: a higher Q lowers the energy needed to detune the ring (narrower line) and gives better channel selectivity, but narrows the optical bandwidth, makes the ring more sensitive to fabrication-induced resonance errors and slows the photon-lifetime-limited response ($\tau_p = Q/\omega \approx 8$ ps for $Q = 10^4$, which is not a limit here but becomes important for $Q > 10^5$).

5. Switch Architectures Based on Micro-Rings

5.1 1×2 and 2×2 switching cells

The simplest cell is an add-drop ring between two crossing waveguides. In the cross state the ring is on resonance and transfers light from one bus to the other; in the bar state the ring is detuned and the signals continue on their original waveguides. Two rings in series (second-order filters) give a flatter pass-band and steeper roll-off, improving crosstalk at the cost of a more complex tuning procedure [2].

5.2 N×N crossbar

In the matrix (crossbar) architecture, N input waveguides and N output waveguides intersect in a grid, with a ring placed at each crossing point. To route input i to output j, the ring at position (i, j) is tuned into resonance. The architecture is strictly non-blocking and requires N^2 rings. The path of a signal passes through a number of off-resonance rings and waveguide crossings that grows linearly with N, and so the insertion loss and crosstalk accumulate. Hitless 4×4 routers based on this principle have been demonstrated on silicon [11], and large silicon switches (up to 32×32 and beyond) have been reported with MZI and ring-assisted designs [12].

5.3 Wavelength-selective and WDM switches

Because an MRR responds only to a particular wavelength, a bank of rings with slightly different radii or tuning can route individual WDM channels from one bus to another. A comb switch can move multiple wavelengths simultaneously by using a cascade of rings, each addressed independently [9]. This allows a single physical path to serve multiple flows and is the basis of ring-based wavelength-routed optical networks-on-chip.

5.4 Broadband switching with rings

To overcome the narrow optical bandwidth of a single ring, ring-assisted MZIs and broadband architectures use rings as phase shifters inside an interferometer, or use a dual-ring structure with a wide FSR. Vlasov et al. demonstrated a wavelength-insensitive silicon switch with micro-ring-based elements for on-chip optical networks [13].

```
flowchart LR
  I1[Input 1] --> R11((Ring 1,1)) --> R12((Ring 1,2)) --> O1[Output 1]
  I2[Input 2] --> R21((Ring 2,1)) --> R22((Ring 2,2)) --> O2[Output 2]
  R11 -. cross .-> R21
  R12 -. cross .-> R22
```

Figure 1. Conceptual 2×2 ring-based crossbar. Rings at the crossing points route a signal between waveguides when tuned to resonance.

6. Performance Metrics

Insertion loss (IL). The power lost in the on-state path, consisting of coupling loss, ring propagation loss, the number of off-resonance rings traversed and waveguide crossings. In a crossbar the worst-case IL grows approximately linearly with N.

Extinction ratio (ER). The ratio between the power at the intended output in the ON state and the OFF state, typically 15–30 dB for a single ring and more for cascaded designs.

Crosstalk. Leakage to unintended ports, caused by the finite roll-off of the Lorentzian line, coupling imperfections, and crossing reflections. Adjacent-channel crosstalk depends on channel spacing relative to the linewidth.

Switching time. Limited by the tuning mechanism (Table 1) and the photon lifetime of the cavity.

Switching energy and power. Thermal switches typically require a few milliwatts of static power per ring, whereas carrier-based switches need femtojoule to picojoule energy per switching event.

Optical bandwidth. A single ring operates in a band of roughly one linewidth for a given state, so aggregate throughput is increased by WDM, not by widening the channel.

Footprint and scalability. A ring occupies 10^2–10^3 μm², orders of magnitude less than an MZI, which supports the dense integration of large switch matrices.

7. Challenges and Mitigation

Fabrication variation. Nanometre-scale variations in waveguide width and thickness shift the resonance by hundreds of picometres to a few nanometres, which exceeds the linewidth. Each ring therefore needs individual calibration and tuning, and this increases power and control complexity. Mitigation: post-fabrication trimming, local heater calibration and feedback control using in-resonator photoconductive monitors [14].

Thermal sensitivity. The ring resonance drifts ~70 pm/K, so a few kelvin of change, caused by neighbouring electronics or by other rings, can detune the device. Mitigation: thermal isolation (trenches, undercut), athermal designs with negative-thermo-optic cladding, and closed-loop wavelength locking [6].

Limited FSR and bandwidth. The FSR (~18 nm in our example) limits the usable number of WDM channels. Smaller rings give a wider FSR but higher bending loss. Mitigation: Vernier configurations, racetrack designs and lower-loss bends.

Insertion loss and crosstalk in large matrices. Accumulated through-port loss and crossings restrict the size of crossbar switches. Mitigation: multi-stage (Clos/Beneš) topologies, low-loss crossings and ring-assisted MZI cells [12].

Control and calibration at scale. An N×N crossbar demands N^2 independent control channels. Practical systems need CMOS drivers, automated algorithms and photonic–electronic co-integration.

Non-volatile and low-power operation. Thermal holding power is significant for circuit-switched systems that rarely change state. Phase-change and ferroelectric materials aim to remove it.

8. Applications

1. **Data-center networks:** reconfigurable optical circuit switches that avoid O-E-O conversion and reduce latency and power [1].
2. **Optical networks-on-chip:** wavelength-routed or circuit-switched fabrics connecting cores or chiplets [11].
3. **Reconfigurable optical add-drop multiplexers (ROADMs):** wavelength-selective routing in metro and access networks.
4. **High-performance computing and AI accelerators:** co-packaged optics for GPU-to-GPU and memory links with high bandwidth density.
5. **Photonic computing and neuromorphic circuits:** ring weight banks in which tunable rings act as programmable weights.
6. **Microwave photonics and sensing:** tunable filters and switching of RF-modulated optical carriers.

9. Future Directions

Several research lines are likely to determine whether MRR switches reach wide deployment. Hybrid integration of high-speed, low-loss Pockels materials (thin-film lithium niobate, barium titanate, electro-optic polymers) with silicon or silicon nitride could provide switching at tens of gigahertz without the free-carrier loss. Non-volatile phase-change cladding offers zero-power holding of states. Machine-learning-assisted calibration and on-chip monitors should reduce the control overhead of large ring arrays. Finally, 3-D and co-packaged integration of the photonic layer with CMOS electronics shortens the interface to the control logic and is expected to lower the energy per bit of the entire system.

10. Conclusion

Micro-ring resonators combine compactness, wavelength selectivity and compatibility with silicon manufacturing, which makes them a leading candidate for the switching elements of future optical interconnects. Their operation is well described by simple analytical transfer functions, and a resonance shift of only about one linewidth is enough for switching. Thermo-optic tuning is simple but slow; carrier-based tuning achieves nanosecond or faster switching with higher loss; emerging materials promise both speed and low power. The main remaining barriers are fabrication sensitivity, thermal control, optical bandwidth and the scalable calibration of thousands of rings. Solutions combining athermal design, automated control and new electro-optic materials are expected to move MRR-based optical switches from laboratory demonstrations toward practical systems.

References

[1] Q. Cheng, M. Bahadori, M. Glick, S. Rumley and K. Bergman, "Recent advances in optical technologies for data centers: a review," Optica, vol. 5, no. 11, pp. 1354–1370, 2018.

[2] B. E. Little, S. T. Chu, H. A. Haus, J. Foresi and J.-P. Laine, "Microring resonator channel dropping filters," J. Lightwave Technol., vol. 15, no. 6, pp. 998–1005, 1997.

[3] W. Bogaerts et al., "Silicon microring resonators," Laser & Photonics Reviews, vol. 6, no. 1, pp. 47–73, 2012.

[4] A. Yariv, "Universal relations for coupling of optical power between microresonators and dielectric waveguides," Electron. Lett., vol. 36, no. 4, pp. 321–322, 2000.

[5] D. G. Rabus, Integrated Ring Resonators: The Compendium, Springer, 2007.

[6] K. Padmaraju and K. Bergman, "Resolving the thermal challenges for silicon microring resonator devices," Nanophotonics, vol. 3, no. 4–5, pp. 269–281, 2014.

[7] R. A. Soref and B. R. Bennett, "Electrooptical effects in silicon," IEEE J. Quantum Electron., vol. 23, no. 1, pp. 123–129, 1987.

[8] Q. Xu, B. Schmidt, S. Pradhan and M. Lipson, "Micrometre-scale silicon electro-optic modulator," Nature, vol. 435, pp. 325–328, 2005.

[9] P. Dong, S. F. Preble and M. Lipson, "All-optical compact silicon comb switch," Opt. Express, vol. 15, no. 15, pp. 9600–9605, 2007.

[10] V. R. Almeida, C. A. Barrios, R. R. Panepucci and M. Lipson, "All-optical control of light on a silicon chip," Nature, vol. 431, pp. 1081–1084, 2004.

[11] N. Sherwood-Droz et al., "Optical 4×4 hitless silicon router for optical networks-on-chip (NoC)," Opt. Express, vol. 16, no. 20, pp. 15915–15922, 2008.

[12] L. Qiao, W. Tang and T. Chu, "32×32 silicon electro-optic switch with built-in monitors and balanced-status units," Sci. Rep., vol. 7, 42306, 2017.

[13] Y. A. Vlasov, W. M. J. Green and F. Xia, "High-throughput silicon nanophotonic wavelength-insensitive switch for on-chip optical networks," Nature Photonics, vol. 2, pp. 242–246, 2008.

Green finance in india: emerging opportunities and future prospects

Aarushi Raj [1], Md Asif Faizi [2]
[1] Student, Department of MBA; [2] Assistant Professor, Department of MBA,
[1,2] Guru Gobind Singh Educational Society's Technical Campus, Kandra, Bokaro, Jharkhand, India

ABSTRACT

Green finance has become an important mechanism for integrating environmental sustainability with economic and financial development. Climate change, pollution, resource depletion and rising energy demand have encouraged governments, financial institutions and investors to direct capital towards environmentally responsible activities. This paper examines the emerging opportunities and future prospects of green finance in India, with a focus on financial instruments, institutional participation, government initiatives and implementation challenges. It adopts a descriptive and exploratory approach based on secondary information from academic studies, regulatory publications and institutional reports. The analysis identifies renewable energy, electric mobility, sustainable infrastructure, agriculture, MSME modernisation, waste management and green technology as key areas of opportunity. Green bonds, green loans, green deposits, ESG investments and blended finance can support the mobilisation of capital. However, high initial costs, limited awareness, greenwashing, inconsistent standards and restricted access for smaller enterprises remain significant barriers. The paper concludes that transparent disclosure, effective regulation, investor education and inclusive financing are essential for strengthening India's green finance ecosystem.

Keywords: Green Finance, Sustainable Development, Green Bonds, ESG Investment, Green Banking, Climate Finance.

1. INTRODUCTION

Economic growth has traditionally been assessed through output, investment, employment and infrastructure development. Yet conventional economic activity can also contribute to pollution, deforestation, greenhouse-gas emissions and the excessive use of natural resources. This has made environmental sustainability an important consideration in economic policy and financial decision-making.

Green finance refers to financial products, services and investments that support environmental protection, climate-change mitigation and adaptation, and the transition towards a low-carbon economy. It includes funding for renewable energy, energy-efficient infrastructure, clean transportation, sustainable agriculture, water conservation and waste management.

Financial institutions influence how capital is allocated through lending, investment and insurance. Banks can incorporate environmental risks into credit decisions, while capital markets can mobilise long-term funds through green bonds and sustainable investment products. India is an important context for this development because it must meet rising energy and infrastructure needs while addressing climate and environmental risks. Green finance can therefore support sustainable growth and create new investment opportunities.

2. RESEARCH PROBLEM AND OBJECTIVES

India requires substantial long-term investment in clean energy, sustainable transport, efficient industries and climate-resilient infrastructure. Conventional financing alone may not meet these needs. Although green finance offers alternative channels, its expansion is constrained by high project costs, limited awareness, measurement difficulties, greenwashing and the limited ability of MSMEs to access specialised finance.

The study aims to: (i) explain the concept and significance of green finance in India; (ii) identify major opportunities and financial instruments; (iii) examine the role of financial institutions and regulatory initiatives; (iv) identify key barriers; and (v) assess future prospects and suggest measures to strengthen the ecosystem.

3. REVIEW OF LITERATURE

Earlier research has examined green bonds, sustainable investment and the relationship between financial markets and environmental performance. Ehlers and Packer (2017) highlighted the importance of certification and standardisation in building confidence in green bonds and ensuring that funds are used for eligible environmental purposes. Zerbib (2019) found that investors' environmental preferences can influence green bond pricing, suggesting that sustainability considerations may affect investment decisions.

Banga (2019) discussed green bonds as a potential source of climate finance for developing countries, while noting challenges in building effective markets. Tang and Zhang (2020) examined market responses to corporate green bond announcements and their implications for shareholders. Flammer (2021) reported that corporate green bond issuance can encourage stronger environmental commitments and improvements in aspects of corporate environmental performance.

Ozili (2022) reviewed global green finance research and identified themes such as climate change, sustainable investment, green banking and financial innovation. More recent reviews, including Woode (2024), emphasise the importance of policy conditions, financial instruments and technological innovation in supporting green growth.

Overall, the literature recognises green finance as a tool for mobilising sustainable investment. However, further work is needed on how Indian regulatory measures, banking participation, MSME access, investor awareness and digital technologies jointly influence the development of the green finance ecosystem.

4. RESEARCH GAP

Existing studies frequently examine individual instruments or environmental outcomes separately. There remains a need for an integrated understanding of the Indian context, particularly in relation to: (a) the affordability and accessibility of green finance for MSMEs and start-ups; (b) the role of banks, NBFCs, insurers and mutual funds; (c) investor awareness and financial literacy; (d) the effectiveness of disclosure and verification requirements; and (e) the use of fintech, artificial intelligence and blockchain to improve monitoring and transparency. This paper addresses these themes through a consolidated review of instruments, policies, opportunities and challenges.

5. RESEARCH METHODOLOGY

The study follows a descriptive and exploratory research design. It is qualitative and conceptual in nature and is based exclusively on secondary data. Sources include academic journal articles, publications of the Reserve Bank of India (RBI) and Securities and Exchange Board of India (SEBI), Government of India policy documents, and reports from international institutions. Descriptive and qualitative content analysis are used to review financial instruments, regulatory initiatives, opportunities and constraints. No primary survey or interviews have been conducted; therefore, the findings are qualitative and are not statistical estimates.

The study focuses on the Indian financial system, including green bonds, green loans, green deposits, ESG investment, green banking, renewable energy and sustainable infrastructure. Its main limitation is dependence on secondary sources and the evolving nature of green finance regulations and reporting practices.

6. MAJOR INSTRUMENTS OF GREEN FINANCE IN INDIA

6.1 Green Bonds and Sustainability-Linked Bonds

Green bonds are debt securities whose proceeds are earmarked for eligible environmental projects, such as renewable energy, clean transport, water management and energy efficiency. Sustainability-linked bonds, in contrast, connect financial terms to the issuer's achievement of specified sustainability targets; their proceeds are not necessarily restricted to particular green projects.

6.2 Green Loans and Green Deposits

Green loans finance activities such as solar installations, energy-efficient machinery and electric-vehicle infrastructure. Green deposits allow eligible financial institutions to raise deposits for allocation to specified green activities, subject to applicable regulatory requirements and reporting.

6.3 ESG Investment, Carbon and Blended Finance

ESG investment considers environmental, social and governance factors alongside financial performance. Carbon finance includes carbon credits and other mechanisms that support emission reduction. Blended finance combines public, development or philanthropic capital with private investment to reduce risks and make sustainable projects more attractive to commercial financiers.

7. EMERGING OPPORTUNITIES FOR GREEN FINANCE IN INDIA

7.1 Renewable Energy

Solar, wind and other renewable energy projects require substantial investment in generation, storage and transmission. Banks, green bond issuers and infrastructure funds can provide long-term capital for equipment, project development and grid-related infrastructure. Financing can also support innovation and employment in the clean-energy sector.

7.2 Electric Mobility and Sustainable Transport

Electric-vehicle manufacturers, battery producers, charging-station developers and fleet operators need capital to expand. Specialised loans from banks and NBFCs, together with investment in battery technology and energy storage, can accelerate the development of cleaner transport systems.

7.3 Green Infrastructure and Urban Development

Green finance can support energy-efficient buildings, sustainable housing, metro systems, water-treatment facilities, waste-management projects and climate-resilient urban drainage. Such investment can reduce environmental costs while meeting the needs of growing cities.

7.4 Sustainable Agriculture and Rural Development

Financing opportunities include solar irrigation pumps, water-efficient systems, climate-resilient farming, organic production and agricultural-waste management. Accessible financial products can help farmers and rural enterprises adopt resource-efficient practices.

7.5 MSMEs and Green Entrepreneurship

MSMEs may need finance to replace outdated machinery, reduce waste and adopt energy-efficient technologies. Credit guarantees, technical assistance and simplified loan procedures can improve access. Green start-ups in recycling, clean technology, sustainable packaging and water management also offer opportunities for venture capital and impact investment.

7.6 Circular Economy, Green Hydrogen and Digital Finance

Recycling, wastewater treatment, material recovery and waste-to-energy projects can convert discarded resources into economic value. Green hydrogen and energy storage are emerging areas requiring investment in research, production and infrastructure. Digital platforms, data analytics and artificial intelligence can support project appraisal, environmental-risk assessment and monitoring. Blockchain may improve traceability in selected applications.

8. ROLE OF FINANCIAL INSTITUTIONS

Commercial banks can offer green loans and deposits, finance renewable energy and sustainable transport, and include environmental risks in credit assessment. NBFCs can serve borrowers and smaller enterprises through specialised products, including electric-vehicle and solar-equipment finance. Capital markets mobilise long-term resources through green and sustainability-linked bonds, while mutual funds, insurers and pension funds can direct institutional investment towards credible sustainable businesses.

Development finance institutions can provide long-term funding and project expertise. In India, institutions such as the Indian Renewable Energy Development Agency (IREDA) support renewable-energy financing. Effective participation requires trained staff, reliable project evaluation and monitoring of environmental outcomes.

9. GOVERNMENT POLICIES AND REGULATORY INITIATIVES

The Government of India introduced the Sovereign Green Bond Framework in 2022 to raise resources for eligible green public expenditure, including renewable energy, clean transportation, energy efficiency, sustainable water management and climate adaptation.

The RBI's Framework for Acceptance of Green Deposits, issued in April 2023 and effective from June 2023, provides guidance for regulated entities accepting green deposits. It emphasises allocation of proceeds, depositor protection, verification and reporting. SEBI has also developed disclosure requirements for green debt securities and sustainability reporting, including the Business Responsibility and Sustainability Reporting framework for listed companies.

India's climate commitments under the Paris Agreement provide a wider policy context for investment in emission reduction and climate resilience. Together, these initiatives help establish an institutional foundation, although consistent implementation and credible reporting remain essential.

10. CHALLENGES AFFECTING GREEN FINANCE IN INDIA

Despite growing policy and market interest, several barriers may limit the expansion of green finance:

High initial investment: Renewable energy, electric mobility and sustainable infrastructure can require substantial upfront capital and long payback periods.

Limited awareness: Retail investors, businesses and smaller borrowers may have limited knowledge of green financial products and their risks.

Greenwashing: Exaggerated or misleading environmental claims can weaken investor confidence and undermine the credibility of green instruments.

Lack of standardisation: Differences in definitions, eligibility criteria, ratings and impact-measurement methods make products difficult to compare.

MSME access constraints: Smaller enterprises may lack collateral, technical expertise and reporting capacity needed to secure specialised finance.

Technology and project risks: New technologies may involve uncertain costs, operating performance and commercial viability.

Limited institutional expertise: Some financial institutions need specialised skills to assess climate risks and verify environmental benefits.

Addressing these barriers requires more than increasing the volume of finance. The quality of project selection, transparency in the use of proceeds and measurement of actual environmental outcomes are equally important.

11. FINDINGS AND DISCUSSION

The review leads to the following principal findings:

Green finance is becoming an important part of India's financial system as regulators, governments and financial institutions increasingly consider environmental sustainability in capital allocation.

Green bonds, green loans, green deposits and ESG investment provide different channels for mobilising capital, but investor confidence depends on clear standards and credible disclosures.

Renewable energy, electric mobility, sustainable infrastructure, agriculture, waste management and green technology represent major opportunities for investment.

Banks, NBFCs, capital markets and institutional investors have a central role in financing the transition and incorporating environmental considerations into financial decisions.

MSMEs and rural enterprises remain an important but underserved segment and require affordable credit, guarantees and technical support.

Digital technologies can improve access, monitoring and transparency, but technology must be supported by reliable data and effective governance.

These findings are based on secondary research and qualitative interpretation. They should not be treated as results of a primary survey or as evidence of a statistical causal relationship between green finance and economic growth.

12. SUGGESTIONS AND RECOMMENDATIONS

Develop consistent definitions, eligibility criteria and reporting standards for green financial products.

Encourage banks to expand specialised green lending and integrate climate and environmental risks into credit assessment.

Improve investor and borrower awareness through financial-literacy programmes and institutional training.

Expand access for MSMEs through credit guarantees, simplified procedures, technical assistance and appropriately designed incentives.

Strengthen independent verification, transparent disclosure and periodic impact assessment to prevent greenwashing.

Promote public-private partnerships and blended finance to reduce risks in projects with high initial costs or long payback periods.

Use digital tools and data analytics to improve project appraisal, monitoring and environmental reporting.

Support research, green entrepreneurship and credible transition plans for carbon-intensive industries.

13. FUTURE PROSPECTS

The future of green finance in India is closely linked to the country's infrastructure needs, energy transition, technological development and sustainability commitments. Green bond markets may expand as public agencies and corporations seek long-term capital. Banks are likely to integrate environmental considerations more systematically into lending and risk management, while ESG information may increasingly influence corporate reporting and institutional investment.

Fintech, artificial intelligence and digital reporting can improve the accessibility and monitoring of sustainable financial products. Climate finance and blended-finance arrangements may attract domestic and international capital to projects that are difficult to fund through conventional channels. Transition finance can also support industries such as steel, cement and power in adopting credible pathways to reduce their environmental impact.

For these prospects to be realised, green finance must become more inclusive. Financing should reach not only large corporations but also MSMEs, farmers, start-ups, households and rural entrepreneurs. Consistent regulation, skilled professionals, reliable environmental data and transparent impact assessment will be critical to long-term market development.

14. CONCLUSION

Green finance provides an important means of connecting financial development with environmental sustainability. This study examined its instruments, emerging opportunities, institutional roles, regulatory initiatives and challenges in India. The analysis indicates that renewable energy, electric mobility, sustainable infrastructure, agriculture, MSME modernisation, waste management and green entrepreneurship offer significant opportunities for investment.

Government and regulatory initiatives, including the Sovereign Green Bond Framework, the RBI's Green Deposit Framework and SEBI's sustainability-related disclosures, provide a foundation for the development of the market. Nevertheless, high initial costs, limited awareness, greenwashing, inconsistent standards and restricted access for smaller enterprises must be addressed.

A stronger ecosystem will require regulatory coordination, innovative financial products, public-private partnerships, investor education, digital monitoring and measurable environmental outcomes. Green finance should therefore be viewed not merely as an alternative investment category, but as a strategic component of India's long-term and inclusive economic development.

REFERENCES

1. Abhilash, P., Shenoy, S. S., Shetty, D. K., & Lobo, L. S. (2023). Green bond as an innovative financial instrument in the Indian financial market: Insights from systematic literature review approach. SAGE Open, 13(2). https://doi.org/10.1177/21582440231178783
2. Banga, J. (2019). The green bond market: A potential source of climate finance for developing countries. Journal of Sustainable Finance & Investment, 9(1), 17–32. https://doi.org/10.1080/20430795.2018.1498617
3. Ehlers, T., & Packer, F. (2017). Green bond finance and certification. BIS Quarterly Review. Bank for International Settlements.
4. Flammer, C. (2021). Corporate green bonds. Journal of Financial Economics, 142(2), 499–516. https://doi.org/10.1016/j.jfineco.2021.01.010
5. Ozili, P. K. (2022). Green finance research around the world: A review of literature. International Journal of Green Economics, 16(1), 56–75. https://doi.org/10.1504/IJGE.2022.125554
6. Reserve Bank of India. (2023). Framework for acceptance of green deposits. https://www.rbi.org.in/
7. Securities and Exchange Board of India. (2023). Revised disclosure requirements for issuance and listing of green debt securities. https://www.sebi.gov.in/
8. Tang, D. Y., & Zhang, Y. (2020). Do shareholders benefit from green bonds? Journal of Corporate Finance, 61, Article 101427. https://doi.org/10.1016/j.jcorpfin.2018.12.001
9. Zerbib, O. D. (2019). The effect of pro-environmental preferences on bond prices: Evidence from green bonds. Journal of Banking & Finance, 98, 39–60. https://doi.org/10.1016/j.jbankfin.2018.10.012

Impact of artificial intelligence on employee productivity and job opportunities

Muskan Kumari [1], Md Asif Faizi [2]
[1] Student, Department of MBA; [2] Assistant Professor, Department of MBA,
[1,2] Guru Gobind Singh Educational Society's Technical Campus, Kandra, Bokaro, Jharkhand, India

ABSTRACT

Artificial Intelligence (AI) is transforming the nature of work by automating routine activities, supporting decision-making and enabling employees to perform tasks more efficiently. At the same time, AI is changing skill requirements, job roles and employment relationships. This paper examines the impact of AI on employee productivity and job opportunities, with particular attention to the skills employees may need and the concerns they may experience. The study adopts a descriptive and exploratory approach based on secondary sources, including academic literature and reports from international organisations. The review indicates that AI can improve productivity by reducing repetitive work, accelerating information processing, supporting data analysis and assisting communication. However, the benefits depend on the quality of implementation, employee training, human oversight and organisational readiness. AI may displace or redesign some tasks while creating demand for roles in data, technology, cybersecurity, AI governance and human–AI collaboration. Major employee concerns include job insecurity, privacy, algorithmic bias, work intensification and inadequate training. The paper concludes that organisations should adopt a human-centred approach that combines technological investment with reskilling, transparent policies and employee participation.

Keywords: Artificial Intelligence, Employee Productivity, Job Opportunities, Automation, Reskilling, Future of Work.

1. INTRODUCTION

Artificial Intelligence has become an important part of modern organisations. AI applications such as machine learning, natural language processing, predictive analytics, intelligent automation and generative AI are increasingly used in areas including human resource management, finance, marketing, customer service, manufacturing and education. These technologies can process large amounts of information, identify patterns, generate content and support routine decisions.

The adoption of AI has created expectations of higher productivity, improved accuracy, faster service delivery and lower operating costs. Employees may use AI tools to prepare reports, analyse data, respond to customer queries, schedule work and generate ideas. However, the same technologies can automate certain tasks and change the demand for particular occupations and skills.

The effect of AI should therefore not be understood only as a choice between job creation and job loss. In many cases, AI changes the tasks performed within a job, alters the level of human judgement required and creates new forms of collaboration between people and machines. The central management challenge is to ensure that productivity gains are accompanied by fair employment practices, skill development and responsible use of technology.

2. STATEMENT OF THE RESEARCH PROBLEM

Organisations are adopting AI to improve efficiency and competitiveness, but employees may not have equal access to the skills and support needed to adapt. Automation can reduce demand for some routine tasks, while new AI-enabled roles may require technical, analytical and interpersonal capabilities. Employees may also be concerned about job security, surveillance, biased decisions and the reliability of AI-generated outputs. The research problem is to understand how AI influences employee productivity and job opportunities, and how organisations can manage the associated benefits and risks.

3. OBJECTIVES OF THE STUDY

- To examine the impact of AI on employee productivity and work performance.
- To identify changing job roles and emerging employment opportunities associated with AI.

- To examine the skills employees may need in an AI-enabled workplace.
- To identify employee concerns and organisational challenges arising from AI adoption.
- To suggest strategies for responsible AI implementation and workforce development.

4. RESEARCH QUESTIONS

The study is guided by three questions: (1) How does AI improve or influence employee productivity? (2) What new skills and job opportunities are emerging as organisations adopt AI? (3) What concerns do employees have about AI, and how can organisations address them?

5. REVIEW OF LITERATURE

Research on automation and technological change suggests that technology affects tasks as well as occupations. Autor (2015) explains that automation can substitute for certain tasks while complementing workers in activities requiring problem-solving, adaptability and interpersonal interaction. This perspective indicates that the impact of AI depends on the tasks within a job rather than simply the job title.

Acemoglu and Restrepo (2020) examine the relationship between automation, new tasks and employment. Their work highlights that automation may displace labour from particular activities, while the creation of new tasks can generate opportunities for workers. The overall employment effect depends on the balance between these forces and on how quickly workers can move into changing roles.

The OECD Employment Outlook 2023 discusses the implications of AI for work, including job quality, skills, algorithmic management and worker well-being. It notes that AI may improve performance and reduce some burdens, but can also create risks related to work intensity, privacy and autonomy. The International Labour Organization (2023) similarly emphasises that generative AI is more likely to augment many occupations than fully automate them, although the effects differ across tasks and groups of workers.

The World Economic Forum's Future of Jobs Report 2025 identifies AI and information-processing technologies among important drivers of labour-market change. It anticipates growth in technology-related skills and roles, while also highlighting the importance of analytical thinking, resilience, flexibility and lifelong learning. These findings reinforce the need for organisations and educational institutions to prepare employees for changing skill requirements.

Recent experimental research by Brynjolfsson, Li and Raymond (2025) on generative AI in customer support found that access to an AI assistant increased productivity, with particularly strong benefits for less experienced workers. The result illustrates how AI can support learning and spread effective practices, although outcomes may vary across occupations and organisational settings.

6. RESEARCH GAP

A substantial body of literature discusses automation, employment and AI-related productivity. However, the effects are not uniform across industries, occupations, employee experience levels or organisational contexts. There is a need for integrated discussion of productivity, emerging job opportunities, changing skill requirements and employee concerns. In particular, organisations need to understand how training, employee participation, transparency and human oversight influence whether AI adoption improves work or increases insecurity. This paper brings these dimensions together in the context of contemporary workplaces.

7. RESEARCH METHODOLOGY

The study uses a descriptive and exploratory research design and is based on secondary data. Information has been drawn from peer-reviewed research and publications of the OECD, International Labour Organization and World Economic Forum. The study uses qualitative content analysis to identify recurring themes relating to productivity, employment, skills and employee concerns. No primary survey or interviews have been conducted; consequently, the paper presents a conceptual review rather than statistical findings from employees. Its conclusions should be interpreted in light of the differences between sectors and the rapid evolution of AI technologies.

8. CONCEPTUAL UNDERSTANDING: AI AND THE WORKPLACE

AI refers to computer systems capable of performing activities commonly associated with human intelligence, such as recognising patterns, understanding language, making predictions and generating content. In workplaces, AI may be used as an assistant that supports employees, as an automation system that performs defined tasks, or as a decision-support tool that provides recommendations. The extent of human review, the consequences of errors and the sensitivity of the data determine the level of oversight required.

Employee productivity may be understood as the quantity and quality of output produced relative to the time and resources used. AI can influence productivity through speed, accuracy, consistency, access to information and support for problem-solving. Yet productivity should not be measured only by the number of tasks completed; quality, customer outcomes, employee well-being and long-term learning also matter.

9. IMPACT OF AI ON EMPLOYEE PRODUCTIVITY

9.1 Automation of Routine Tasks

AI-enabled systems can automate repetitive and rule-based activities such as data entry, document classification, invoice processing, appointment scheduling and basic customer enquiries. By reducing time spent on routine work, employees may devote more attention to complex tasks, customer relationships and problem-solving. Automation nevertheless requires monitoring because incorrect data, system failures or unusual cases may produce errors.

9.2 Faster Information Processing and Decision Support

AI tools can search, summarise and analyse large datasets more quickly than manual processes. In finance, they may assist with transaction monitoring and forecasting; in HR, they may support workforce analytics; and in marketing, they may help analyse customer trends. These tools can improve the speed of decisions, but managers remain responsible for checking assumptions, context and accuracy.

9.3 Improved Communication and Service Delivery

Generative AI and chatbots can help prepare drafts, answer common questions, translate text and provide round-the-clock support. Employees can use these tools to improve response times and consistency. Human involvement remains important for sensitive, complex or emotionally demanding interactions.

9.4 Learning and Support for Less Experienced Employees

AI assistants can provide examples, explanations and step-by-step guidance. This may help new employees learn processes more quickly and enable experienced employees to focus on complex work. Organisations should ensure that employees understand the limits of AI and do not rely on unverified outputs.

9.5 Risks to Productivity

AI does not automatically increase productivity. Poorly selected tools, inaccurate outputs, weak data quality, inadequate training and fragmented workflows can create additional work. Excessive monitoring or unrealistic performance targets may also increase stress. Productivity improvements therefore depend on thoughtful implementation, employee training and evaluation of both output quality and employee well-being.

10. CHANGING JOB OPPORTUNITIES

AI is likely to change the composition of employment rather than affect all jobs in the same way. Some routine tasks may be reduced, while new activities emerge around developing, operating, securing and governing AI systems. Examples of growing or evolving roles include:

AI and machine-learning specialists who develop, test and maintain AI applications.

Data analysts and data engineers who organise information and generate insights.

Cybersecurity and information-governance professionals who protect systems and data.

AI auditors, risk analysts and governance specialists who assess reliability, fairness and compliance.

Automation specialists and business-process analysts who redesign workflows.

Human–AI collaboration roles in customer service, education, healthcare, finance and other professional fields.

AI may also increase the value of occupations that require empathy, negotiation, leadership, creativity, complex judgement and physical work in unpredictable environments. At the same time, some roles may be redesigned, and workers performing highly automatable tasks may face greater transition pressures. The availability of new jobs does not guarantee that displaced workers can move into them without suitable training and support.

11. SKILLS REQUIRED IN AN AI-ENABLED WORKPLACE

Employees will need a combination of technical, cognitive and human skills. Important capabilities include:

AI literacy: understanding what AI can and cannot do, and recognising limitations and risks.

Data literacy: interpreting data, checking evidence and communicating insights.

Digital and cybersecurity awareness: using digital tools safely and protecting confidential information.

Critical thinking: verifying AI-generated content, identifying errors and making sound judgements.

Creativity and problem-solving: applying knowledge to unfamiliar situations and developing new solutions.

Communication, teamwork and adaptability: collaborating across functions and learning continuously.

Reskilling and upskilling should be treated as continuous organisational responsibilities rather than one-time training activities.

12. EMPLOYEE CONCERNS ABOUT ARTIFICIAL INTELLIGENCE

12.1 Job Insecurity and Displacement

Employees may fear that automation will reduce the need for their roles or limit future career opportunities. Uncertainty can affect morale, trust and willingness to adopt new systems. Organisations should communicate honestly about the purpose of AI adoption and provide transition and retraining support wherever possible.

12.2 Privacy and Workplace Surveillance

AI-based monitoring tools may analyse productivity, attendance, communications or work patterns. If used without clear limits, such systems can make employees feel constantly monitored. Organisations should explain what data is collected, why it is needed, who can access it and how long it is retained.

12.3 Bias and Unfair Decisions

AI systems trained on biased or incomplete data may reproduce unfair patterns. This is particularly concerning in recruitment, promotion, performance evaluation and workforce allocation. Human review, regular testing, explainability and accessible appeal mechanisms are important safeguards.

12.4 Accuracy, Accountability and Overdependence

AI-generated information can be incorrect, outdated or misleading. Employees may also become overly dependent on automated recommendations. Clear accountability, verification procedures and human oversight are essential, especially where decisions affect people, finances, safety or legal rights.

12.5 Unequal Access to Training

Employees in smaller organisations, informal work and lower-paid roles may have fewer opportunities to develop AI-related skills. Unequal access can widen existing labour-market inequalities. Training should be accessible, practical and designed for different levels of digital experience.

13. FINDINGS AND DISCUSSION

The review of literature and institutional reports supports the following broad findings:

AI can improve employee productivity by automating routine tasks, accelerating information processing and supporting communication and decision-making.

Productivity gains depend on appropriate workflows, reliable data, employee skills and human review; technology alone does not guarantee better performance.

AI is likely to transform many jobs by changing task content, while creating demand for roles in data, AI development, cybersecurity, automation and governance.

AI literacy, analytical thinking, adaptability, creativity and communication are important capabilities for employees.

Job insecurity, privacy, bias, inaccurate outputs and unequal access to training are major employee concerns.

A human-centred implementation strategy can help organisations balance efficiency with fairness, accountability and employee well-being.

These findings are qualitative conclusions from secondary sources. They are not results of a primary employee survey and should not be interpreted as applying equally to every occupation or industry.

14. SUGGESTIONS AND RECOMMENDATIONS

Conduct a task-level impact assessment before introducing AI, identifying which activities can be automated and which require human judgement.

Provide continuous AI literacy, digital skills and role-specific reskilling programmes for employees.

Involve employees and their representatives in the design, testing and evaluation of workplace AI systems.

Establish clear policies on data privacy, security, acceptable use and accountability for AI-assisted decisions.

Maintain human oversight in high-impact decisions such as recruitment, promotion, discipline and termination.

Evaluate AI systems for accuracy, bias, accessibility and unintended consequences at regular intervals.

Use productivity measures that consider quality, customer outcomes and employee well-being, not only speed or volume.

Support employees whose tasks are significantly affected through redeployment, career guidance and transition assistance.

15. FUTURE PROSPECTS

AI is expected to become more integrated into everyday organisational processes. Rather than functioning only as a separate technology, it may increasingly operate as a workplace assistant embedded in office software, customer platforms, production systems and analytical tools. This can support faster workflows and more informed decisions, provided that organisations maintain appropriate safeguards.

The future labour market is likely to place greater value on the ability to work effectively with AI. Employees who combine domain knowledge with AI literacy, data interpretation and human-centred skills may be better positioned to adapt. Educational institutions and employers will need to update curricula and training programmes to reflect changing occupational requirements.

The distribution of benefits will depend on policy and management choices. If AI is introduced mainly to reduce labour costs, some employees may experience displacement and reduced job quality. If it is used to augment human capabilities, improve working conditions and create new products and services, it may support broader productivity and employment opportunities. Responsible governance, worker consultation and lifelong learning will therefore shape the social and economic outcomes of AI adoption.

16. CONCLUSION

Artificial Intelligence is reshaping employee productivity and job opportunities by automating routine activities, supporting decisions and changing the skills required in the workplace. AI can help employees work faster, analyse information and improve service delivery. It can also transform existing occupations and create demand for new roles in technology, data, cybersecurity, automation and AI governance.

However, these benefits are accompanied by concerns relating to job insecurity, privacy, bias, accuracy, overdependence and unequal access to training. The impact of AI is therefore not predetermined; it depends on how organisations choose to design, implement and govern these technologies.

The study concludes that organisations should adopt a human-centred approach that combines technological innovation with employee participation, continuous reskilling, transparent policies and human oversight. AI should be treated as a tool to enhance human capabilities rather than as a substitute for responsible management. Such an approach can help organisations improve productivity while preparing employees for changing job opportunities.

REFERENCES

1. Acemoglu, D., & Restrepo, P. (2020). Robots and jobs: Evidence from US labor markets. Journal of Political Economy, 128(6), 2188–2244. https://doi.org/10.1086/705716
2. Autor, D. H. (2015). Why are there still so many jobs? The history and future of workplace automation. Journal of Economic Perspectives, 29(3), 3–30. https://doi.org/10.1257/jep.29.3.3
3. Brynjolfsson, E., Li, D., & Raymond, L. R. (2025). Generative AI at work. The Quarterly Journal of Economics, 140(2), 889–942. https://doi.org/10.1093/qje/qjae044
4. International Labour Organization. (2023). Generative AI and jobs: A global analysis of potential effects on job quantity and quality. ILO.
5. OECD. (2023). OECD employment outlook 2023: Artificial intelligence and the labour market. OECD Publishing. https://doi.org/10.1787/08785bba-en
6. World Economic Forum. (2025). The future of jobs report 2025. World Economic Forum.

Adoption of Digital Supply Chain Management and Its Effect on Operational Efficiency in the Indian Automobile Industry

Shalu Kumari, Student, MBA Department
Mahavir Prasad*, Assistant Professor, MBA Department
Guru Gobind Singh Educational Society's Technical Campus (GGSESTC), Bokaro, Jharkhand, India

ABSTRACT

This paper examines how far digital supply chain management (DSCM) has entered the Indian automobile industry and what it appears to have done for operational efficiency. It uses only secondary data for FY2021-22 to FY2025-26: SIAM and ACMA statistics, PIB and NCAER logistics reports, PLI policy documents, Mahindra & Mahindra and Tata Motors disclosures, and peer-reviewed literature. Sources are read through a documentary evidence-mapping method that groups claims under cost, speed, inventory, quality and flexibility and grades their strength. Three findings emerge. First, the industry has reached record volumes but remains exposed to shocks such as the semiconductor shortage and the 2025 rare earth magnet restrictions. Second, clear digital efforts are visible in large OEMs (end-to-end visibility, control towers, live scenario planning), while evidence on Tier-2 and Tier-3 suppliers is thin. Third, gains are reported mostly in qualitative terms, so the data support a link between digital adoption and efficiency but do not prove cause and effect. A standard public measure of digital adoption in auto supply chains is recommended.

Keywords: digital supply chain; Industry 4.0; operational efficiency; Indian automobile industry; secondary data analysis; supply chain visibility

1. INTRODUCTION

India is among the world's largest automobile markets. In FY2025-26 the industry recorded its highest annual sales in seven years: about 28.27 million domestic units (+10.4%), 4.64 million passenger vehicle wholesales, over 21.7 million two-wheelers, and exports up about 24% to roughly 6.65 million units [S1–S5]. The auto component industry reached Rs 7.60 lakh crore (USD 85.9 billion), up 12.7% [S8][S9]. Yet volume growth does not mean a healthy supply chain. The 2021-22 semiconductor shortage cut passenger vehicle output by around 100,000 units, about 4% [S12]. After China tightened rare earth magnet export controls in April 2025, automakers holding only four to six weeks of stock faced possible production cuts, with India importing nearly 90% of such magnets from China [S12][S13][S14]. The Nexperia takeover later raised new chip worries [S15]. The firms did not lack the ability to build vehicles; they lacked early, shared and reliable information about the deeper tiers of the chain. That is the gap DSCM is meant to close.

DSCM here means using connected data and digital tools (real-time tracking, shared planning platforms, control towers, analytics and AI, automation, machine-vision quality inspection, traceability records, digital twins) to plan, run and monitor flows of materials, information and money between a vehicle maker, suppliers, logistics partners and customers. **Operational efficiency** is read through five dimensions: cost, delivery speed and reliability, inventory control, quality and flexibility. Most research on digital supply chains relies on manager surveys or single-firm cases, often from developed economies; Indian studies list technologies and barriers but rarely bring together the public data published by industry bodies, regulators and companies. This paper asks: (RQ1) what do public data show about the scale, growth and supply risks of the industry; (RQ2) what evidence shows adoption of DSCM by vehicle makers and suppliers; (RQ3) what outcomes do sources report for cost, speed, inventory, quality and flexibility; and (RQ4) how strong is this evidence and where are the gaps?

2. OBJECTIVES OF THE STUDY

(1) To describe the size, growth and supply chain risks of the Indian automobile and auto component industry; (2) to identify digital supply chain practices reported by vehicle makers and suppliers; (3) to map these practices to the five efficiency dimensions and record reported outcomes; (4) to grade the strength of the evidence and point out data gaps; and (5) to suggest actions for firms and policy makers.

3. LITERATURE REVIEW AND THEORETICAL BACKGROUND

Supply chain management began as a drive to move goods at the lowest cost, then added enterprise software, web links with partners and, more recently, connected machines and real-time data. Büyüközkan and Göçer (2018) describe the digital supply chain as one built on smart technologies and shared information to create value for all partners. Key technologies are IoT devices, big data analytics (whose benefit comes from using data in decisions, not collecting it; Kache & Seuring, 2017), AI and machine learning, cloud platforms, blockchain traceability, robotic process automation and digital twins. Evidence from several countries suggests digital tools help performance, but not automatically: Ivanov, Dolgui and Sokolov (2019) argue they help firms see disruption early and limit the ripple effect, and Wamba et al. (2017) found analytics improves performance mainly where organisational capabilities are strong. Reported barriers are software and sensor costs, poor data quality, legacy systems, skills shortages, data security worries and weak buyer–supplier trust, which fall hardest on small suppliers and can create a digital divide. Three theories frame the study: the Technology–Organisation–Environment framework (Tornatzky & Fleischer, 1990), in which adoption depends on the technology, the firm's resources and the environment (customers, competitors, policy); the resource-based view (Barney, 1991), which explains unequal benefits from the same tools; and dynamic capabilities, which stress continual sensing and reshaping of resources.

4. CONCEPTUAL FRAMEWORK AND PROPOSITIONS

With no survey data, the study examines five propositions against documentary evidence rather than testing hypotheses. DSCM adoption is the starting point, visibility and integration the route, operational efficiency the outcome, and shocks and policy the environment that pushes firms to adopt. **P1:** firms reporting wider use of digital tools also report better operational outcomes. **P2:** visibility is the main route through which digital tools improve outcomes. **P3:** closer digital integration with suppliers goes with better delivery and inventory outcomes. **P4:** adoption is uneven, with large OEMs ahead of smaller suppliers. **P5:** shocks and policy support increase the motive to adopt.

5. RESEARCH METHODOLOGY

The study is descriptive and exploratory, with a documentary secondary-data design; no questionnaire or interview was used. It combines descriptive trend analysis of published statistics, content analysis of company and policy documents, and a comparative reading of two large OEMs, Tata Motors and Mahindra & Mahindra, chosen because both publish detailed annual reports and lead across segments. Sources are industry bodies (SIAM, ACMA), government and research bodies (PIB, NCAER for DPIIT), company disclosures, rating-agency and trade press (CRISIL, CareEdge, ICRA as reported by media; Economic Times) and academic literature. The period is FY2021-22 to FY2025-26, with monthly releases to mid-2026 for context. A document was included only if it came from an identifiable body, was dated, reported figures or specific practices, and related to the Indian automobile or component sector; opinion pieces without data were excluded, and official sources were preferred over press summaries. Each statement was coded for the digital practice, the efficiency dimension and the type of evidence, then compared across sources (triangulation). Evidence was graded A (measured figure from an official or industry body), B (measured or specific company claim) or C (general descriptive claim without figures). Reliability was supported by recording source, date and basis for every data point, and validity by triangulating at least two source types per major finding. Company claims are treated as evidence of what firms say they do, not as independent proof of results. All data are public and no personal information was used.

6. INDUSTRY CONTEXT FROM SECONDARY DATA

SIAM figures show FY2025-26 closing at record levels after an uneven start: sales dipped in the first half and rose 16.7% in the second, helped by GST rate reforms and repo rate cuts [S2][S3]. Such swings matter because parts planning must follow demand at short notice. SIAM warned that West Asia conflicts could disturb demand, supply chains and production [S2][S3]. ACMA (over 1,100 members, over 90% of organised-sector turnover) reports the component indicators in Table 1 [S7].

Table 1. Indian vehicle and auto component indicators, FY2025-26

Indicator	FY2025-26	Change on FY2024-25	Source
Domestic vehicle sales (all)	28,265,519 units	+10.4%	[S2]

Passenger vehicles / two-wheelers (wholesales)	4,643,439 / 21,705,974 units	+7.9% / +10.7%	[S1]
Commercial vehicles, domestic	1.08 million units	+12.6%	[S4]
Total vehicle exports	about 6.65 million units	+24%	[S3]
Component industry turnover	USD 85.9 bn (Rs 7.60 lakh crore)	+12.7% (FY25: USD 80.2 bn)	[S8][S9][S26]
Supply to OEMs / aftermarket		+16.3% / +9%	[S8]
Component exports / imports	USD 24 bn / USD 25.4 bn	+5% / +13%	[S6][S8]
Component trade balance	deficit about USD 1.37 bn	from surplus about USD 0.5 bn	[S8][S26]
EV components, share of domestic OEM supplies	4.6%	n.r.	[S8]

Two points stand out. Supply to OEMs grew faster than the aftermarket, so OEM dependence on suppliers is rising along with volumes; and imports grew faster than exports, moving the trade balance to deficit, which raises the need for shipment visibility (the authors' reading, not ACMA's). ACMA's website notes that the Automotive Mission Plan 2026 targeted USD 200 billion component turnover by 2026, against USD 85.9 billion reported [S6].

Logistics cost. Figures of 13–14% of GDP were long quoted, but PIB, drawing on an NCAER study for DPIIT, now reports 7.97% of GDP (9.09% of non-services output; about Rs 24.01 lakh crore) in 2023-24, saying older figures rested on partial data [S10][S11]. This is economy-wide, not auto-specific, but it suggests remaining inefficiency lies less in freight rates than in delays, poor coordination and lack of information (an interpretation). The National Logistics Policy 2022, PM Gati Shakti and ULIP are the main government responses [S10][S11].

Disruptions and lessons. The semiconductor shortage (2021-22) showed that dependence on a few chip sources can cut output sharply [S12]. China's rare earth controls (April 2025) showed that deep-tier concentration (over 90% of magnet processing is in China) is hard to see without mapping suppliers beyond Tier-1 [S12][S13][S14]; commentary added that stable pandemic-era supply had made firms complacent and just-in-time had left them exposed [S14]. The Nexperia takeover (late 2025) showed the value of early warning and clear communication [S15], and West Asia and shipping-route tensions (2026) show that route and port risk needs live monitoring [S2][S16]. By May 2026 leading passenger vehicle makers were treating disruption as a lasting structural challenge [S16].

Policy support. The PLI scheme (outlay Rs 25,938 crore; 82 approved applicants) had attracted Rs 44,326 crore of investment and 67,820 jobs by March 2026, with Rs 2,386.36 crore disbursed [S22]; nearly all payouts went to five vehicle makers [S23]. The sources describe incentives in terms of sales, investment and domestic value addition, not digital adoption, so policy supports capacity and localisation but does not yet reward or measure digital supply chain capability.

7. EVIDENCE OF DSCM ADOPTION IN COMPANY DISCLOSURES

Mahindra & Mahindra. The 2023-24 report lists a supply chain control tower, AI-based analytics, collaborative robots and a digital stockyard system managing 500+ tractor varieties [S18]. The 2025-26 report describes a company-wide programme on a unified digital backbone: end-to-end digital visibility, digitised promise management, real-time execution monitoring (said to improve adherence to customer commitments, though no figure is given) and fully digitised planning with live scenario planning over 200+ constraints [S17]. Press reports add raised inventory buffers, supplier diversification and real-time disruption monitoring [S16].

Tata Motors. Tata Motors reports AI across product development, manufacturing, quality control, after-sales and workforce management, including AI vision systems, condition-based monitoring and digital quality-control platforms for earlier defect detection [S19]; its CTO says machine learning is applied to supply chain and retail optimisation [S20]. Its supply chain strategy rests on localisation, supplier diversification, value-chain collaboration

and selective buffers for critical parts [S16], and a five-year digital platform with TCS includes Scope 3 emissions tracking [S24]. Chairman N. Chandrasekaran has argued that chains built for efficiency must now be built for resilience [S21].

Pattern and gaps. Both companies invest in digital visibility and planning while keeping physical safeguards (buffers, local sourcing, multiple suppliers): digital tools are used alongside, not instead of, traditional safeguards, and efficiency is being redefined to include resilience. No source gives an industry-wide measure of how many OEMs or suppliers use specific digital tools; Tier-2 and Tier-3 adoption is almost invisible, and ACMA's release gives no adoption rates [S27]. The lack of a common measure is itself a finding.

8. FINDINGS

Table 2 maps reported digital practices to the efficiency dimensions and grades the evidence; Table 3 gives the verdict on each proposition.

Table 2. Evidence map by efficiency dimension

Dimension	Digital practice reported	Outcome reported	Grade
Cost	National logistics integration; OEM planning analytics	Logistics cost 7.97% of GDP; no firm-level saving quantified	A (national), C (firm)
Speed and reliability	End-to-end visibility, promise management, real-time monitoring (Mahindra)	Better adherence to commitments; no figure	B
Inventory	Real-time monitoring; live scenario planning	Buffers raised for critical parts, so inventory rose	B
Quality	AI vision, condition-based monitoring (Tata Motors)	Earlier defect detection; no figure	C
Flexibility	Scenario planning over 200+ constraints; supplier diversification	Greater ability to re-plan and shift sourcing; no measure	B

Table 3. Assessment of the propositions

Proposition	Verdict	Reason
P1: wider digital use goes with better outcomes	Partly supported	Firms report better outcomes, but figures are missing and claims are self-reported
P2: visibility is the main route	Supported	Most concrete digital claims concern visibility, monitoring and planning; shocks arose from poor deep-tier visibility
P3: supplier integration improves delivery and inventory	Not enough evidence	Sources describe collaboration and diversification, not digital supplier links or their results
P4: adoption is uneven	Indirectly supported	Detailed disclosures come only from large OEMs; PLI payouts concentrated in five makers [S16][S23]
P5: shocks and policy raise motive to adopt	Supported	Strategies changed after chip, magnet and shipping shocks; policy stresses logistics efficiency [S10][S16]

Summary. (1) The industry is large and growing, which raises the cost of any supply chain failure. (2) Shocks exposed weak visibility into deeper supplier tiers, especially for chips and rare earth magnets. (3) Large OEMs report clear digital efforts in visibility, planning, analytics and quality control. (4) Reported gains are mostly qualitative, so the link between adoption and efficiency is plausible but not proven. (5) Evidence below Tier-1 is very limited and no standard public measure of digital adoption exists.

9. DISCUSSION

The evidence agrees with earlier research: the costliest episodes arose from risks several tiers down that firms could not see in time, and leading OEMs have responded by building visibility and live planning (Ivanov et al., 2019). Mahindra's digitised planning linked to execution monitoring suggests an organisation-wide capability rather than a single tool, consistent with Wamba et al. (2017) and the resource-based view. The TOE framework explains the uneven picture: shocks and policy push all firms, but money, skills and bargaining power let large OEMs move faster, as the concentration of PLI payouts hints (though it is not a direct measure of digital adoption). Efficiency is

also being redefined: after 2021 and 2025, leading firms hold more buffer stock for critical parts while using digital information to decide where a buffer is worth its cost. Finally, with logistics cost under 8% of GDP, reliability, coordination and information become the main sources of avoidable loss, which digital tools target, though the data cannot yet show how much they achieve.

10. IMPLICATIONS

Vehicle makers should map suppliers beyond Tier-1, link visibility tools to planning so alerts lead to decisions, place buffers where they matter most, and publish measured results such as on-time delivery and inventory days. **Component suppliers** should start with low-cost order and stock sharing with the OEM, then move to analytics, using OEM-led training and shared platforms. **Industry bodies** (SIAM, ACMA) could add a regular digital readiness measure to performance reviews. **Policy makers** could add digital capability to incentive criteria or reporting, support MSME suppliers and keep ULIP open to auto supply chains. **Researchers** can use the evidence map to design surveys and cases that measure results with figures.

11. LIMITATIONS AND FUTURE RESEARCH

The study relies on public documents, so it shows patterns, not cause and effect. Company reports present firms favourably and seldom give figures for gains; only two OEMs are examined and no supplier-level documents were available; and some figures come from trade press repeating industry-body releases, so they should be checked against original SIAM, ACMA and government publications. Future work can combine this evidence map with manager surveys, build matched OEM–supplier cases, track a panel of listed component firms using ratios such as inventory days, and study EV supply chains.

12. CONCLUSION

The Indian automobile industry is bigger than ever, but its supply chains have been tested hard. Public data show strong growth, rising import dependence and repeated shocks. Large makers respond with digital visibility, planning and analytics alongside buffers and wider sourcing. Visibility is the clearest route by which digital tools support efficiency, but the evidence does not yet prove how large the gains are and says little about smaller suppliers. A standard public measure of digital adoption, clearer firm reporting of results and targeted support for smaller suppliers would strengthen future research and help the whole chain.

Declarations: The authors declare no conflict of interest. No specific funding was received [edit if applicable]. All data come from public secondary sources listed below; no human participants were involved. Acknowledgements: [add, or delete this line].

REFERENCES

1. Barney, J. (1991). Firm resources and sustained competitive advantage. Journal of Management, 17(1), 99–120.
2. Büyüközkan, G., & Göçer, F. (2018). Digital supply chain: Literature review and a proposed framework for future research. Computers in Industry, 97, 157–177.
3. Ivanov, D., Dolgui, A., & Sokolov, B. (2019). The impact of digital technology and Industry 4.0 on the ripple effect and supply chain risk analytics. International Journal of Production Research, 57(3), 829–846.
4. Kache, F., & Seuring, S. (2017). Challenges and opportunities of digital information at the intersection of Big Data Analytics and supply chain management. International Journal of Operations & Production Management, 37(1), 10–36.
5. Tornatzky, L. G., & Fleischer, M. (1990). The processes of technological innovation. Lexington Books.
6. Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. Journal of Business Research, 70, 356–365.

Supply Chain Adaptiveness and Marketing Responsiveness in the FMCG Sector

Rajnish Kumar, MBA 2nd Semester
Vikash Kumar Jain, Assistant Professor, MBA Department
Department of Management Studies, Guru Gobind Singh Educational Society's Technical Campus (GGSESTC), Chas, Bokaro, Jharkhand, India

Abstract

Purpose: The fast-moving consumer goods (FMCG) sector operates under thin margins, short product life cycles, volatile demand and intense retail-channel competition. Firms in this sector are therefore compelled to align what they can supply with what the market wants, at speed. This paper examines how supply chain adaptiveness (SCA) and marketing responsiveness (MR) relate to each other and jointly shape firm performance in the FMCG sector.

Design/methodology/approach: The study is based exclusively on secondary data: peer-reviewed literature, industry reports (IBEF, NielsenIQ, CII, McKinsey, Deloitte), company annual reports and sustainability disclosures, and documented episodes of disruption such as GST implementation (2017), the Maggi ban (2015), the COVID-19 lockdowns (2020-21), input-cost inflation (2021-23) and the rise of quick commerce. A structured content-analysis approach and a comparative multiple-episode analysis are used to develop a conceptual framework and a set of research propositions.

Findings: The evidence indicates that (i) supply chain adaptiveness is a precondition for credible marketing responsiveness, because promotions, pack innovations and channel shifts fail when the supply network cannot execute them; (ii) marketing responsiveness, in turn, supplies the demand-sensing information that makes adaptiveness purposeful rather than merely reactive; (iii) the two capabilities are complementary, and firms that develop them together appear to recover faster from shocks; and (iv) data analytics, distribution-network flexibility, and cross-functional integration are the main enablers.

Originality/value: The paper integrates two literatures that are usually treated separately (supply chain management and market orientation) in the Indian FMCG context, and offers a testable framework for subsequent primary research.

Keywords: Supply chain adaptiveness; marketing responsiveness; FMCG; supply chain agility; market orientation; demand sensing; India

1. Introduction

Fast-moving consumer goods (FMCG), also called consumer packaged goods (CPG), are products that are sold quickly and at relatively low cost: packaged foods, beverages, personal care, home care, and over-the-counter health products. Margins per unit are small, and profitability depends on volume, shelf availability, and speed of replenishment. A stock-out of a toothpaste or a biscuit brand on a retailer's shelf is typically not a lost sale of that brand alone but a switch to a competitor, since brand loyalty in low-involvement categories is fragile (Fisher, 1997).

India's FMCG sector is among the country's largest, with a distinctive structure: a very large number of small traditional outlets (kirana) serving most demand; a substantial rural base; a growing modern-trade and e-commerce presence; and, recently, quick-commerce platforms promising delivery in minutes. IBEF and NielsenIQ describe rural markets as contributing a significant share of revenue and urban consumption as increasingly digital (the latest releases should be consulted for current figures).

Repeated shocks characterise this environment. GST (2017) forced redesign of warehouses and distribution points; the 2015 Maggi ban showed how quickly a dominant brand can lose shelf presence; the COVID-19 lockdowns (2020-21) shut factories and transport while pushing demand towards hygiene, immunity and pantry products; and inflation in 2021-23 raised palm oil, packaging and logistics costs, forcing changes in pack sizes, prices and promotions. Each episode tested both operations and marketing's ability to read and respond to the market.

Two concepts explain why some firms cope better. **Supply chain adaptiveness** is the ability of a supply chain to adjust its structure, processes and partnerships to longer-term shifts in markets, technologies and regulations (Lee, 2004). **Marketing responsiveness** is the extent to which a firm acts on market intelligence by modifying products, prices, promotions and distribution (Kohli & Jaworski, 1990; Jaworski & Kohli, 1993). The two meet where a market signal must become a delivered product on a shelf, yet the literature studies them in silos: operations scholars focus on agility, resilience and flexibility, marketing scholars on market orientation and customer responsiveness.

This paper attempts to bridge the two. Using secondary data, it asks how supply chain adaptiveness and marketing

responsiveness interact in the FMCG sector, and what this implies for managers and researchers.

1.1 Statement of the problem

FMCG firms invest heavily in demand generation (advertising, trade promotions, launches) and logistics (depots, distributors, transport, technology), but these are planned and evaluated by different functions with different metrics. Promotions run without stock, launches are delayed by sourcing constraints, and supply plans are disrupted by demand swings marketing could have anticipated. The problem is a lack of integrated understanding of how adaptive supply chains and responsive marketing reinforce one another.

1.2 Significance of the study

The study consolidates dispersed secondary evidence into one framework; it is set in an emerging market, where infrastructure variability and fragmented retail make adaptiveness harder; and it yields propositions testable through surveys or structural equation modelling.

2. Review of Literature

2.1 Supply chain adaptiveness, agility, flexibility and resilience

The vocabulary in this field overlaps and requires clarification. Flexibility is the capacity to change within a predefined range at low cost and time. Agility is the capability to respond rapidly to short-term changes in demand or supply. Resilience is the ability to anticipate, withstand and recover from disruption (Christopher & Peck, 2004; Sheffi & Rice, 2005). Adaptability or adaptiveness concerns the longer-term adjustment of the network itself.

Lee (2004) argued in his well-known "Triple-A" framework that the best supply chains are not only agile but also adaptable and aligned. Agility responds to short-term fluctuations; adaptability adjusts the supply chain design to structural market shifts; alignment ensures that the interests of all partners are consistent with those of the firm. This is a useful lens for FMCG, where both daily replenishment volatility and structural change, such as the rise of e-commerce, coexist.

Fisher (1997) distinguished functional products, which have predictable demand and call for efficient supply chains, from innovative products, which have unpredictable demand and call for responsive ones. Most FMCG staples are functional in nature, but promotions, seasonal products, new variants and regional preferences introduce innovative-product characteristics. FMCG firms therefore generally need a hybrid supply chain, efficient at the core and responsive at the edge.

Swafford, Ghosh and Murthy (2006) showed that supply chain flexibility is built through procurement, manufacturing and distribution/logistics flexibility, supported by IT integration. Braunscheidel and Suresh (2009) found that internal and external integration and a learning culture contribute to agility. Gligor, Esmark and Holcomb (2015) and later studies showed that agility can improve firm performance, particularly in turbulent environments. Ivanov and Dolgui (2020) extended this to the viability of supply networks during COVID-19, stressing redundancy, flexibility and digital visibility.

2.2 Marketing responsiveness and market orientation

Kohli and Jaworski (1990) defined market orientation as the organisation-wide generation of market intelligence, dissemination of that intelligence across departments, and responsiveness to it. Narver and Slater (1990) described it as consisting of customer orientation, competitor orientation and inter-functional coordination. Jaworski and Kohli (1993) provided evidence that market orientation is related to business performance, with the strength of the relationship moderated by market turbulence and competitive intensity.

Day (1994) positioned market sensing and customer linking as core capabilities of a market-driven organisation, and noted that such capabilities are only valuable when the firm can deliver on what it senses. This observation anticipates the argument of the present paper: marketing responsiveness has a built-in dependency on operations. Responsiveness that stops at the planning document does not constitute responsiveness.

In FMCG, marketing responsiveness appears in several forms: rapid pack-size and price-pack architecture adjustments (for example, low-unit-price sachets for rural and low-income consumers), regional variant launches, dynamic trade promotions, rapid digital campaign changes, and channel-specific assortments for modern trade, e-commerce and quick commerce.

2.3 The interface between supply chain and marketing

The marketing–supply chain interface is a long-recognised source of friction: marketing prefers variety, availability and promotional freedom; operations prefer stability, long runs and forecast accuracy. Research on demand-supply

integration and sales and operations planning (S&OP) suggests that cross-functional processes, shared forecasts and shared metrics reduce this friction. Chopra and Sodhi (2004) argued that risk mitigation must be tied to commercial priorities. Demand sensing, which uses near-real-time point-of-sale and market signals to adjust short-term forecasts, is highlighted in consulting publications by McKinsey and Deloitte.

2.4 Research gap

The literature reviewed shows that (a) supply chain agility and adaptiveness are positively associated with performance, (b) market orientation and responsiveness are positively associated with performance, but (c) relatively few studies examine these two capabilities together, particularly in an emerging-market FMCG context with a fragmented retail structure. A further gap is the paucity of studies that use documented real-world disruption episodes as comparative cases. This study responds to these gaps.

3. Objectives of the Study

9. To conceptualise supply chain adaptiveness and marketing responsiveness in the context of the FMCG sector.
10. To examine, using secondary evidence, the relationship between supply chain adaptiveness and marketing responsiveness.
11. To analyse how major disruption and change episodes in the Indian FMCG sector have exposed or rewarded these capabilities.
12. To identify the key enablers and barriers of integrated adaptiveness and responsiveness.
13. To propose a conceptual framework and testable propositions for future empirical research.

4. Research Methodology

4.1 Research design

The study follows a descriptive and exploratory design based entirely on secondary data. It combines (a) a structured literature review and (b) a comparative analysis of documented episodes from the FMCG sector.

4.2 Sources of secondary data

Secondary data were drawn from academic journals (Journal of Marketing, HBR, Journal of Operations Management, IJPR); industry and government reports (IBEF, NielsenIQ, Kantar, CII, FICCI); consulting publications (McKinsey, Deloitte, BCG, Bain, EY); company annual, integrated and sustainability reports (HUL, ITC, Nestlé India, Dabur, Britannia, Marico); and reputable business press on documented events.

4.3 Method of analysis

Sources were screened for relevance to supply chain adaptiveness, agility, flexibility or resilience, and to marketing responsiveness, market orientation, or demand sensing, with an emphasis on the FMCG sector. The material was then coded thematically under four headings: sensing, decision and design, execution, and learning. Four disruption or change episodes were compared on how marketing and supply chain responses were coordinated. Because the data are secondary and drawn from public documents, the analysis is interpretive; no statistical hypothesis testing is claimed.

4.4 Scope and delimitations

The paper focuses on the Indian FMCG sector with reference to global literature. Firm examples are illustrative of publicly reported practice and are not a statistical sample. Figures cited should be verified against the latest editions of the cited reports before citation in a final publication.

5. Conceptual Framework and Propositions

5.1 Defining the constructs

Supply chain adaptiveness (SCA) is defined here as the capability of the supply network to sense change, redesign its sourcing, manufacturing and distribution configuration, and realign partner relationships in order to remain effective as market conditions change. It has four dimensions:

- Sourcing flexibility: multiple or alternative suppliers, flexible contracts, and substitutable inputs.

- Manufacturing flexibility: multi-product lines, short changeovers, and the use of contract and third-party manufacturing.
- Distribution flexibility: multi-tier distributor networks, multi-channel fulfilment, and re-routable logistics.
- Information visibility and integration: shared data on inventory, sales and capacity across partners.

Marketing responsiveness (MR) is defined as the speed and appropriateness with which a firm modifies its marketing mix in response to market intelligence. Its dimensions are:

- Product responsiveness: variants, pack sizes and reformulations.
- Price responsiveness: price-pack architecture and promotional pricing.
- Promotion responsiveness: communication and trade promotions.
- Channel responsiveness: channel-specific assortments and fulfilment.

5.2 Linking the constructs

The framework rests on a sense-decide-deliver-learn cycle. Marketing is the primary sensing and deciding function (consumer insights, trade feedback, competitor moves), while the supply chain is the primary delivering function. Learning closes the loop, as sales and fulfilment data feed back into forecasts and planning. Adaptiveness without market sensing is directionless; responsiveness without adaptive supply is unfulfillable.

Figure 1. Conceptual framework (textual representation)

Enablers (data analytics, S&OP/cross-functional integration, partner trust, digital visibility) → Supply Chain Adaptiveness ⇄ Marketing Responsiveness → Customer outcomes (availability, satisfaction, loyalty) → Firm performance (market share, growth, margin). Moderators: environmental turbulence, channel complexity, regulatory change.

5.3 Research propositions

- **P1:** Supply chain adaptiveness is positively related to marketing responsiveness, because adaptive supply networks widen the range of marketing actions that can be executed credibly.
- **P2:** Market intelligence generation and dissemination are positively related to supply chain adaptiveness, because purposeful network redesign depends on accurate reading of the market.
- **P3:** The joint presence of high supply chain adaptiveness and high marketing responsiveness is associated with higher firm performance than either capability alone (a complementarity effect).
- **P4:** Cross-functional integration (for example, effective S&OP) strengthens the relationship between supply chain adaptiveness and marketing responsiveness.
- **P5:** The positive effects of both capabilities on performance are stronger under higher environmental turbulence and channel complexity.

6. Analysis and Discussion of Secondary Evidence

This section compares four documented episodes. In each, the key question is whether the firm's supply chain could adapt and whether its marketing could respond, and whether the two were coordinated. The descriptions rely on widely reported public information and company disclosures and should be cross-checked against primary documents before final submission.

6.1 Episode 1: GST implementation (2017)

The unified indirect tax removed state-level tax boundaries that had shaped warehouse locations. Many FMCG firms consolidated depots into fewer, larger hubs and re-optimised distribution, illustrating structural adaptiveness (network redesign) as distinct from short-term agility. For marketing, the challenges were transitional: distributor destocking, price-list changes and re-labelling. Firms that coordinated price and pack changes with distributor inventory plans saw smoother transitions. The case supports P1 and P4.

6.2 Episode 2: Maggi regulatory ban (2015)

In 2015 Nestlé India withdrew Maggi noodles following regulatory action and later relaunched the brand. On the supply side, the firm had to recall and destroy stock, re-test products and rebuild distribution; on the marketing side, rapid communication, trust rebuilding and a staged relaunch were needed while competitors used the gap to gain shelf space. Marketing responsiveness without availability is incomplete, and restored availability without marketing recovery does not rebuild share quickly. The case supports P3 (complementarity).

6.3 Episode 3: COVID-19 lockdowns (2020-21)

Lockdowns produced a dual shock: constrained manufacturing and logistics, and a sharp demand shift towards hygiene, health, immunity and pantry staples. Publicly reported responses by large FMCG firms included reallocating plant capacity to soaps, sanitisers and disinfectants, prioritising high-velocity SKUs, launching hygiene and immunity products quickly, moving to e-commerce and direct-to-consumer channels, and using contract manufacturers and local transporters to sustain supply. Marketing sensed the demand shift, the supply chain reconfigured capacity and SKU focus, and sales data fed back into planning; firms with granular distributor and retailer data (information visibility) redirected supply more precisely. The episode supports P1, P2 and P5.

6.4 Episode 4: Input-cost inflation and channel shift (2021-24)

Rising prices of palm oil, crude-linked packaging and freight forced calibrated price increases, grammage reductions at held price points, smaller unit-price packs for price-sensitive and rural markets, and a focus on premium products with more resilient margins. Meanwhile quick commerce and e-commerce grew rapidly, requiring distinct pack assortments, faster replenishment and dark-store logistics, while kirana stores still accounted for the bulk of sales. Firms therefore ran parallel supply chains for general trade, modern trade and e-commerce/quick commerce, showing that channel responsiveness (marketing) is feasible only with distribution flexibility (supply chain). The episode supports P1, P4 and P5.

6.5 Comparative summary

Episode	Nature of change	Supply chain adaptiveness evidence	Marketing responsiveness evidence	Key lesson
GST (2017)	Structural, regulatory	Depot consolidation, network redesign	Price-list and pack transition management	Structural adaptiveness enlarges marketing options
Maggi ban (2015)	Brand and regulatory shock	Recall, re-testing, rebuilt distribution	Trust rebuilding, staged relaunch	Availability and communication must recover together
COVID-19 (2020-21)	Demand and supply shock	Capacity reallocation, SKU prioritisation, partner use	New hygiene/immunity products, channel shift	Sensing plus visibility enables fast reconfiguration
Cost inflation and quick commerce (2021-24)	Cost and channel shift	Parallel channel supply chains	Price-pack architecture, channel-specific assortments	Channel responsiveness requires distribution flexibility

Table 2. Comparative summary of episodes

6.6 Enablers identified

Across the episodes, five enablers recur in the secondary material:

1. **Data analytics and demand sensing:** point-of-sale, distributor-management-system and e-commerce data used to adjust forecasts quickly.
2. **Cross-functional planning (S&OP):** shared forecasts and decisions across marketing, sales, supply and finance.
3. **Distribution network design:** multi-tier distribution, rural reach programmes and flexible fulfilment options.
4. **Partner ecosystems:** contract manufacturers, third-party logistics and distributors with aligned incentives.
5. **Leadership and culture:** a willingness to reprioritise quickly and to treat availability as a marketing metric.

6.7 Barriers identified

Barriers include siloed functional metrics (for example, marketing measured on volume and operations on cost), long-lead procurement of imported inputs, limited visibility into deep-tier suppliers and small retailers, infrastructure variability, and legacy IT systems. For smaller FMCG firms, limited capital for technology and for redundancy also constrains adaptiveness.

7. Key Findings

Based on the secondary evidence reviewed, the paper arrives at the following findings:

1. Marketing responsiveness depends on supply chain adaptiveness for execution. Several marketing actions (promotions, launches, channel shifts) are feasible only within the limits of what the network can deliver.
2. Supply chain adaptiveness depends on marketing intelligence for direction. Without sensing, network adjustments risk being misaligned with demand.
3. The two capabilities appear to be complementary. Episodes in which marketing and supply chain responses were coordinated show smoother recovery than those in which one function led and the other lagged.
4. The value of both capabilities rises with turbulence and channel complexity, which are defining features of the Indian FMCG market.
5. Hybrid supply chain design (efficient at the core, responsive at the edge) is the dominant practical pattern for managing the trade-off between cost and speed (consistent with Fisher, 1997).
6. Information visibility and cross-functional planning are the most frequently cited enablers; siloed metrics are the most frequently cited barrier.

8. Implications

8.1 Managerial implications

- **Integrate planning:** institutionalise S&OP with marketing as a full participant and with shared metrics such as on-shelf availability, forecast accuracy for promoted items and order-fill rate.
- **Design for channels:** build channel-specific supply chains and assortments for general trade, modern trade and quick commerce rather than a single network serving all.
- **Invest in sensing:** use retailer-level and distributor-level data to detect changes early, and link sensing directly to production and distribution decisions.
- **Build optionality:** maintain alternative suppliers and contract manufacturing options for critical inputs, and decide where redundancy is worth its cost.
- **Treat availability as a brand asset:** include supply performance in brand health reviews.
- **Support small and rural channels:** programmes that digitise and support small retailers strengthen both the information flow and the reach of the network.

8.2 Theoretical implications

The paper suggests that market orientation theory should treat deliverability as a boundary condition of responsiveness, and that supply chain adaptiveness research should treat market sensing as an antecedent of purposeful adaptation. An integrated view may improve the explanatory power of both literatures.

8.3 Policy implications

Policy measures that improve logistics infrastructure, reduce regulatory uncertainty, and support digital adoption among small retailers can improve the adaptive capacity of the whole sector, not just of large firms.

9. Limitations and Scope for Future Research

The study is based on secondary data and is therefore constrained by the availability, accuracy and objectivity of published sources. Company disclosures may present a favourable picture, and episode-based analysis cannot establish causality. The framework is conceptual and has not been statistically validated. Industry figures change frequently and should be updated at the time of publication.

Future research could (a) test P1 to P5 through a cross-sectional survey of FMCG managers using structural equation modelling, with validated scales for supply chain agility/adaptiveness and for market orientation/responsiveness; (b) conduct in-depth case studies of individual firms; (c) compare large multinational firms with regional and small FMCG firms; (d) examine the roles of artificial intelligence, predictive analytics and sustainability in adaptiveness; and (e) study the effects of quick commerce on traditional distribution.

10. Conclusion

Supply chain adaptiveness and marketing responsiveness are two sides of one commercial capability: sensing what customers want and delivering it reliably as conditions change. In Indian FMCG, where demand is volatile, channels fragmented and shocks recurrent, neither is sufficient alone. Evidence from GST, the Maggi ban, COVID-19 and the cost-inflation and quick-commerce period shows that firms perform best when sensing, decision-making and execution are connected through shared data, cross-functional planning and flexible networks. The framework and propositions offer a basis for empirical testing and practical direction for firms seeking to be both adaptive and responsive.

References

1. Braunscheidel, M. J., & Suresh, N. C. (2009). The organizational antecedents of a firm's supply chain agility for risk mitigation and response. Journal of Operations Management, 27(2), 119-140.
2. Chopra, S., & Sodhi, M. S. (2004). Managing risk to avoid supply-chain breakdown. MIT Sloan Management Review, 46(1), 53-61.
3. Christopher, M., & Peck, H. (2004). Building the resilient supply chain. The International Journal of Logistics Management, 15(2), 1-14.
4. Day, G. S. (1994). The capabilities of market-driven organizations. Journal of Marketing, 58(4), 37-52.
5. Fisher, M. L. (1997). What is the right supply chain for your product? Harvard Business Review, 75(2), 105-116.
6. Gligor, D. M., Esmark, C. L., & Holcomb, M. C. (2015). Performance outcomes of supply chain agility: When should you be agile? Journal of Operations Management, 33-34, 71-82.
7. Ivanov, D., & Dolgui, A. (2020). Viability of intertwined supply networks: Extending the supply chain resilience angles towards survivability. International Journal of Production Research, 58(10), 2904-2915.
8. Jaworski, B. J., & Kohli, A. K. (1993). Market orientation: Antecedents and consequences. Journal of Marketing, 57(3), 53-70.
9. Kohli, A. K., & Jaworski, B. J. (1990). Market orientation: The construct, research propositions, and managerial implications. Journal of Marketing, 54(2), 1-18.
10. Lee, H. L. (2004). The triple-A supply chain. Harvard Business Review, 82(10), 102-112.
11. Narver, J. C., & Slater, S. F. (1990). The effect of a market orientation on business profitability. Journal of Marketing, 54(4), 20-35.
12. Sheffi, Y., & Rice, J. B. (2005). A supply chain view of the resilient enterprise. MIT Sloan Management Review, 47(1), 41-48.
13. Swafford, P. M., Ghosh, S., & Murthy, N. (2006). The antecedents of supply chain agility of a firm: Scale development and model testing. Journal of Operations Management, 24(2), 170-188.

Enhanced Pneumatic Control System For Precise, Jerk-Free Actuation Under Bi-Directional And Vibrating Load Conditions

Akhilesh Anand[a] , Atul Kumar[a] ,Kumar Ashish [b]

Department of Mechanical Engineering, GGSESTC Bokaro

Abstract

Conventional pneumatic actuators are attractive in industrial automation because compressed air is readily available and pneumatic equipment is generally straightforward to integrate. However, precise and stable positioning can become difficult when an actuator operates against a load that changes direction or contains vibration. This paper presents a conceptual enhanced pneumatic control system intended to improve the controllability of linear and angular actuators under such conditions. The proposed arrangement uses a modified mechanically operated spool valve and a charge–switch–control sequence. In the neutral/charging position, compressed air is supplied to both sides of a double-acting actuator piston. For a selected direction, the valve changes the flow path and traps air on the side that is to be controlled, while pressure is released from the opposite side. Small movements of the spool are then intended to regulate the release of compressed air and hence the actuator stroke rate. The concept aims to provide smoother, approximately uniform, low-jerk actuation with reduced bouncing, without requiring electronic control or major changes to the existing actuator and compressed-air source. The supplied design material includes a conceptual spool arrangement and initial CAD views, but does not report measured performance or experimental validation. Accordingly, the system is presented as a development concept and a basis for prototype testing rather than as a demonstrated performance improvement.

Keywords: Pneumatic control; modified spool valve; double-acting actuator; bi-directional load; pressure control; jerk-free actuation; vibrating load.

1. Introduction

Pneumatic systems are widely used where clean, comparatively simple and economical actuation is required. Their use can be less satisfactory in applications that demand controlled displacement while the actuator is loaded. Changes in external force, load reversal and vibration may disturb actuator motion and make it difficult to obtain a stable stroke. Hydraulic systems can provide strong force and controlled motion, but some low-load applications may not require the force capability or level of hydraulic equipment. The design concept considered here seeks an intermediate option: retaining a pneumatic power source while improving the mechanical control of actuator movement.

The proposed system focuses on control of a double-acting actuator in both directions. Its central idea is to pressurize both sides of the piston before motion, select the intended direction, and then meter the release of pressure so that the actuator moves toward the lower-pressure side. The original presentation identifies smooth displacement, reduced jerk and bounce, operation under bi-directional and vibrating loads, and compatibility with existing pneumatic networks as its principal objectives.

2. Aim and design objectives

The aim is to develop a mechanically controlled pneumatic arrangement that offers more stable displacement control in loaded operation. The objectives stated in the design presentation are:

- to improve control of stroke displacement rate in linear and angular actuators;
- to make the actuator position more resistant to external loads in either direction by pressurizing both piston chambers;
- to obtain approximately uniform, jerk-free actuation with negligible bouncing under the intended working-load range;
- to reduce the delay associated with changing the spool position during loaded operation;
- to achieve sturdy, controlled actuation without electronic control; and
- to limit modifications to the existing actuator and pressure source, principally changing the control spool and adding the required pipe connections.

3. Proposed system configuration

The proposed arrangement consists of a compressed-air reservoir or supply, a modified multi-port spool valve, connecting pipes and a double-acting linear actuator. The valve is the principal modified component. The presentation describes the spool visually as a 7-port × 5-position arrangement and states that, when considered according to the referenced spool-design conventions, its functional nature is described as a 4 × 5 design. This distinction should be finalized in a detailed valve schematic and port-identification drawing before manufacture.

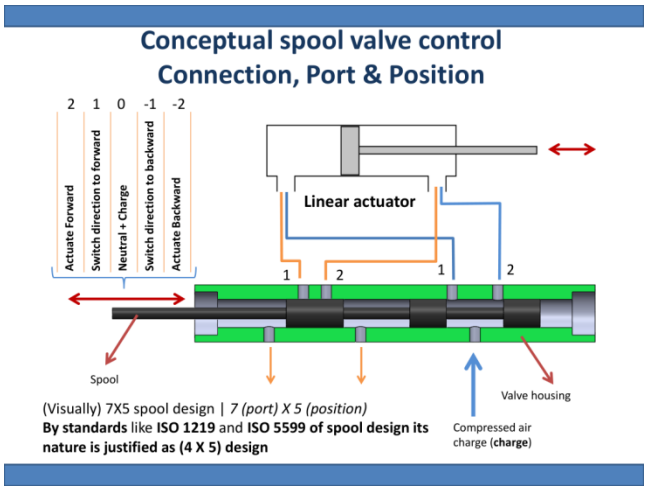


Fig. 1. Conceptual spool-valve port and position arrangement reproduced from the supplied design presentation (slide/page 4).

The initial CAD views show a longitudinal valve housing containing a shaped spool, multiple air connections and a manual lever-and-spring operating mechanism. These views communicate the proposed mechanical layout; they do not, by themselves, establish dimensions, tolerances, sealing performance, pressure rating or manufacturability.

4. Working principle: charge–switch–control

The operating sequence is based on three stages. The working principle is illustrated in Fig. 2.

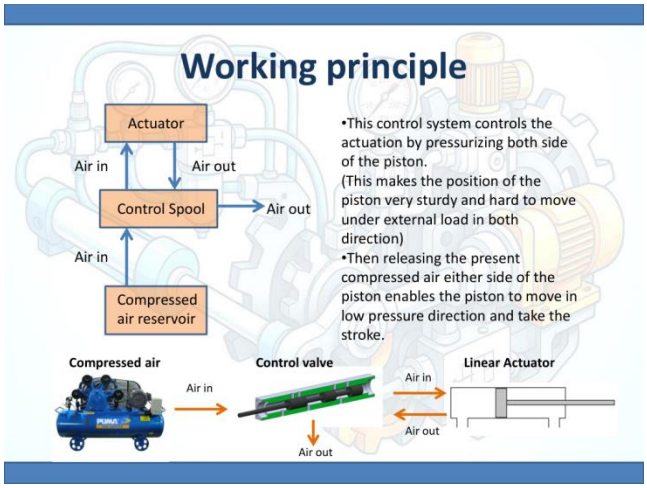


Fig. 2. Functional flow concept: compressed-air supply, control spool and linear actuator (slide/page 3).

4.1 Neutral/charging stage

At the neutral position, the valve supplies compressed air to both sides of the actuator piston. The intention is to establish pressure in both chambers before initiating movement. With pressure acting on both sides, the piston is expected to resist unwanted displacement caused by an external load. The design presentation refers to this condition as "Neutral + Charge"; no pressure values or chamber-balance calculations are provided.

4.2 Direction-selection stage

To move the actuator forward or backward, the spool is shifted to the corresponding switching position. The flow path on the non-selected side is closed so that compressed air is trapped in that chamber, while the selected stroke path is prepared for controlled pressure release. The same operating logic is used for both directions; only the selected side and direction of movement change.

4.3 Controlled actuation stage

After direction selection, the spool is moved by a small amount in the desired direction. This movement is intended to release compressed air progressively and allow the piston to travel toward the lower-pressure side. The degree of spool movement is therefore proposed as the mechanical means of controlling actuator speed and displacement. Reversing the spool command enables the corresponding backward or forward operation. The concept relies on pressure management and mechanical spool positioning rather than electronic feedback.

5. Expected benefits and potential applications

The design presentation identifies the following expected benefits: reduced jerk and bouncing in linear and angular actuators; improved bi-directional control under loaded conditions, including loads with vibration; controlled actuation with little operating delay; and compatibility with existing pneumatic circuits because the actuator and pressure source need not be substantially modified. The proposal may therefore be relevant to industrial shop-floor pneumatic networks where a conventional actuator is already installed but smoother motion is required.

These benefits should be interpreted as design objectives, not verified results. The presentation does not provide test data, measured positioning error, speed variation, settling time, bounce amplitude, energy consumption or a comparison with a conventional directional-control valve. Suitability for a particular application will depend on the actuator size, supply pressure, load range, air compressibility, leakage, valve geometry and operating frequency.

6. Prototype development and validation plan

The supplied presentation includes initial CAD views of the spool assembly and its manual operating mechanism (Fig. 3). The next development stage should convert the conceptual arrangement into a dimensioned engineering design, define the port map and spool positions, and manufacture a prototype suitable for controlled testing.

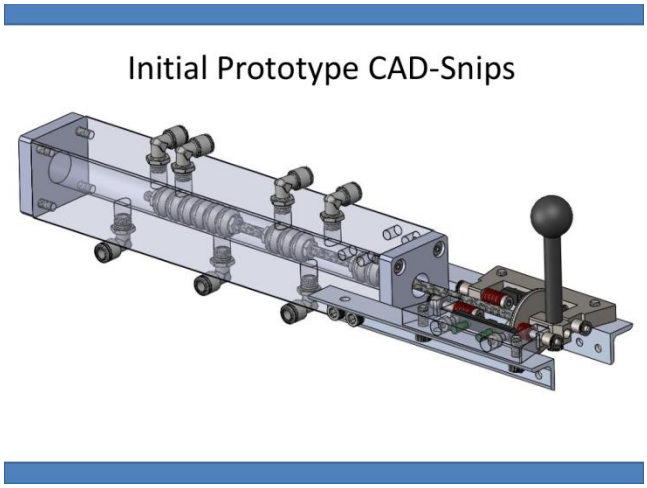


Fig. 3. Initial prototype CAD views provided in the presentation (slide/page 9).

and load. Tests should include unloaded motion, steady opposing load, load reversal, and an imposed vibrating or fluctuating load. Forward and backward strokes should be evaluated separately. Relevant measurements include actuator displacement versus time, velocity variation, overshoot or rebound, response delay, repeatability, leakage and air consumption. Tests should be repeated across a stated operating-pressure and load range. Instrumentation and test conditions should be documented so that claims of precise, jerk-free or bounce-free motion can be quantified.

The design also requires engineering checks of spool clearances, sealing, friction, wear, pressure losses, safe exhaust routing and the possibility of pressure being trapped in either actuator chamber. The valve's port numbering and the

relationship between the visually illustrated 7 × 5 layout and the stated 4 × 5 functional designation should be resolved in the final circuit diagram. These checks are necessary before industrial deployment.

7. Conclusion

An enhanced pneumatic control concept has been presented for improving the motion control of linear and angular actuators under bi-directional and vibrating load conditions. Its main feature is a modified mechanically operated spool valve that follows a charge–switch–control sequence: both actuator chambers are pressurized in neutral, a direction is selected by changing the flow paths, and controlled pressure release is used to produce the required stroke. The approach is intended to provide smoother and more stable actuation while retaining compatibility with existing pneumatic equipment and avoiding electronic control. The supplied material establishes the operating concept and provides preliminary CAD representations, but it does not include experimental results. Prototype manufacture, detailed valve specification and comparative testing are therefore essential next steps to establish the system's actual accuracy, dynamic response, reliability and practical operating range.

Variation of Brake Thermal Efficiency of I C Engine with increase in Ethanol in Ethanol-Petrol Blends

Amit Kumar[1]
Gulam Hussin[2]
Dr. Bikash Ghoshal[3]
[1& 2] Student of 3rd Year, Mechanical Engineering, GGSESTC, Kandra, Chas, Bokaro-827013, Jharkhand
[3]Professor, Mechanical Engineering, GGSESTC, Kandra, Chas, Bokaro-827013, Jharkhand

Abstract: Petrol mixed with ethanol is emerging as a key fuel and research area with the aim of reducing pollution and establishing ethanol as dependable fuel for Spark Ignition (SI) engines (Petrol Engine). Renewable sources of energy are most favored in view of severe climate change affecting the whole mankind. Also, there is uncertainty in the steady fuel supply due to war. This paper analyses the effect of ethanol-petrol blends on brake thermal efficiency. Taguchi's experimental design method is adopted for experimentation. The analysis of variance (ANOVA) is done with the help of Minitab 22.

Key words: Ethanol; efficiency; Taguchi

1. Introduction

Energy requirement is increasing day by day due to population explosion and the associated demand for increasing number of vehicles and also for industrial expansion. As the fossil fuel reserves continue to deplete, alternative energy sources have become inevitable. Renewable sources of energy are considered the best to ensure sustainability. Ethanol (C_2 H_5 OH) has emerged as a promising alternate fuel for Spark Ignition (SI) engines as it is produced from renewable sources of plant materials rich in sugars and starch. Since ethanol is derived from plant-based sources that absorb CO_2 during their growth, its overall carbon footprint is negligible, contributing to climate change mitigation. Researchers have been actively working to reduce dependency on fossil fuels and explore cost-effective, eco-friendly alternatives. Several studies have introduced various fuels that can serve as viable substitutes while minimizing environmental impact. Jitendra Kumar et al. studied the impact of ethanol-blended gasoline on a spark ignition (SI) engine. Ethanol blends of 5%, 10%, 15%, and 20% were tested on a four-cylinder, four-stroke petrol engine. The authors found that ethanol blending reduced exhaust emissions, increased mechanical efficiency but specific fuel consumption increased due to ethanol's lower calorific value [1]. Achinta Sarkar et al. studied the performance and emission characteristics of a spark ignition (SI) engine running on ethanol-gasoline blends. Authors reported that ethanol increased brake torque, indicated power and volumetric efficiency but also raised brake-specific fuel consumption (BSFC) [2]. Gangadharan Nair et al. investigated the performance of a two-stroke spark-ignition (SI) engine using petrol ethanol blends (4% to 32%) under various load conditions. Results showed the optimal blend for maximum brake thermal efficiency (34.6%) and minimum SFC (0.26 kg/kW-hr) was 24% ethanol [3]. Experimentations and analysis by above researchers have their own limitations and need further improvements for sustainable applications.

This study aims at selection of optimum ethanol-petrol blend ratio for maximum brake thermal efficiency. The settings of three input variables of load (Kg), RPM and % of Ethanol:Petrol is determined by L9 Orthogonal arrays of Taguchi. The parameters are varied at three levels each for getting response of brake thermal efficiency.

2. Experimental Details

2.1 Experimental Setup

A "maruti suzuki 800 cc" four cylinder, four stoke multi-cylinder spark engine (SI) as shown in Figure 1 is used for experimentation.

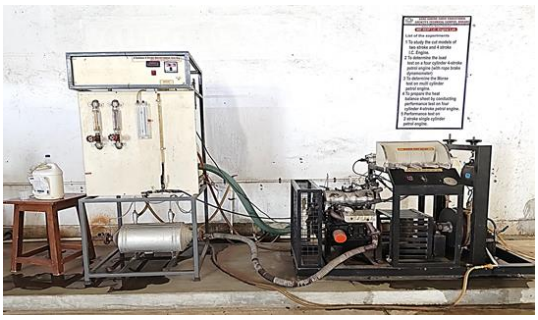


Figure 1 Experimental setup

2.2 Process Parameters

The machining parameters and their levels are shown in Table 1. The settings of three process parameters of load (kg), RPM and ethanol : petrol % are determined by orthogonal arrays of Taguchi. The parameters are judiciously selected based on their influencing effect on brake thermal efficiency. Parameters are varied at three levels each for getting responses of thermal efficiency.

Table 1 Input variable and their levels

Load (Kg)	RPM	Ethanol: Petrol %
12	1000	5
15	1300	10
18	1600	15

2.3 Experimental Design

Taguchi's experimental design method is adopted for experimentation. The settings of three input variables of load (Kg), RPM and % of Ethanol: Petrol is determined by L9 Orthogonal arrays of Taguchi as shown in Table 2. The parameters are varied at three levels each for getting response of efficiency. Total nine experiments are to be performed, and efficiency will be calculated based on standard formula. After experimentation, parameters will be analyzed by the S/N ratio with the concept of larger the better for optimizing efficiency. Essentially, traditional experimental design procedures are too complicated and not easy to use. A large number of experimental works have to be carried out when the number of process parameters increase. To solve this problem, the Taguchi method uses a special design of orthogonal arrays to study the entire parameter space with only a small number of experiments. Taguchi methods have been widely utilized in engineering analysis and consist of a plan of experiments with the objective of acquiring data in a controlled way, in order to obtain information about the behavior of a given process. The greatest advantage of this method is the saving of effort in conducting experiments; saving experimental time, reducing the cost, and discovering significant factors quickly. Taguchi's robust design method is a powerful tool for the design of a high-quality system. In addition to the S/N ratio, a statistical analysis of variance (ANOVA) can be employed to indicate the impact of process parameters on metal removal rate values. The steps applied for Taguchi optimization in this study are as follows:

a) Select Taguchi orthogonal array
b) Conduct Experiments
c) Brake thermal efficiency calculation
d) Analyze results; (Signal-to-noise ratio)
e) Predict optimum performance

Table 2 Taguchi design of L9 experiments

Expt. No	Load (Kg)	RPM	Ethanol: Petrol (%)	Thermal Efficiency (%)
1	12	1000	5	
2	12	1300	10	
3	12	1600	15	
4	15	1000	10	
5	15	1300	15	
6	15	1600	5	
7	18	1000	15	

8	18	1300	5	
9	18	1600	10	

3. Experimentation

3.1 Fuel blends preparation :

3.3.1 BLEND RATIO :

E5 (5% Ethanol 95% Petrol)
E10 (10% Ethanol 90% Petrol)
E15 (15% Ethanol 85% Petrol)

3.2 Experimental procedure

Table 3 Observation table for blend test

Observation	Ethanol petrol Blend ratio	RPM	Load (kg)	BHP	Mass flow rate per kg (30cc)	n_{Bth}
1	5:95	1000	12	2.58	$3.21*10^{-4}$	13.93
2	5:95	1600	15	5.15	$5.04*10^{-4}$	17.14
3	5:95	1300	18	5.02	$4.59*10^{-4}$	18.94
4	10:90	1300	12	3.35	$3.84*10^{-4}$	15.42
5	10:90	1000	15	3.22	$3.36*10^{-4}$	16.94
6	10:90	1600	18	6.18	$5.41*10^{-4}$	20.17
7	15:85	1600	12	4.13	$4.68*10^{-4}$	15.89
8	15:85	1300	15	4.19	$4.04*10^{-4}$	18.70
9	15:85	1000	18	3.87	$3.78*10^{-4}$	18.44

4. Analysis of the S/N Ratio

Taguchi method stresses the importance of studying the response variation using the signal – to – noise (S/N) ratio, resulting in minimization of quality characteristic variation due to uncontrollable parameter. The efficiency was considered as the quality characteristic with the concept of "the larger-the-better". The efficiency values measured from the experiments and their corresponding S/N ratio values are shown in Table 4.5. Regardless of the category of the performance characteristics, a greater S/N value corresponds to a better performance. Therefore, the optimal levels of the performance parameters are the level with the greatest S/N value. Based on the analysis of the S/N ratio, the optimal performance for efficiency obtained at 18 kg load (level 3), 1600 RPM (level 3), and 10 % Ethanol: Petrol blends (level 2). A3 B3 C2 is the optimum efficiency. Fig. 4.5.2 shows the effect of the process parameters on efficiency values.

Load	RPM	% of Ethanol	Efficiency (%)	S/N Ratio
12	1000	5	13.93	22.88
12	1300	10	15.42	23.76

TABLE 4 Experimental combinations of parameters, efficiency and S/N ratio

12	1600	15	15.89	24.02
15	1000	10	16.94	24.58
15	1300	15	18.7	25.44
15	1600	5	17.74	24.98
18	1000	15	18.44	25.32
18	1300	5	18.94	25.55
18	1600	10	21.17	26.51

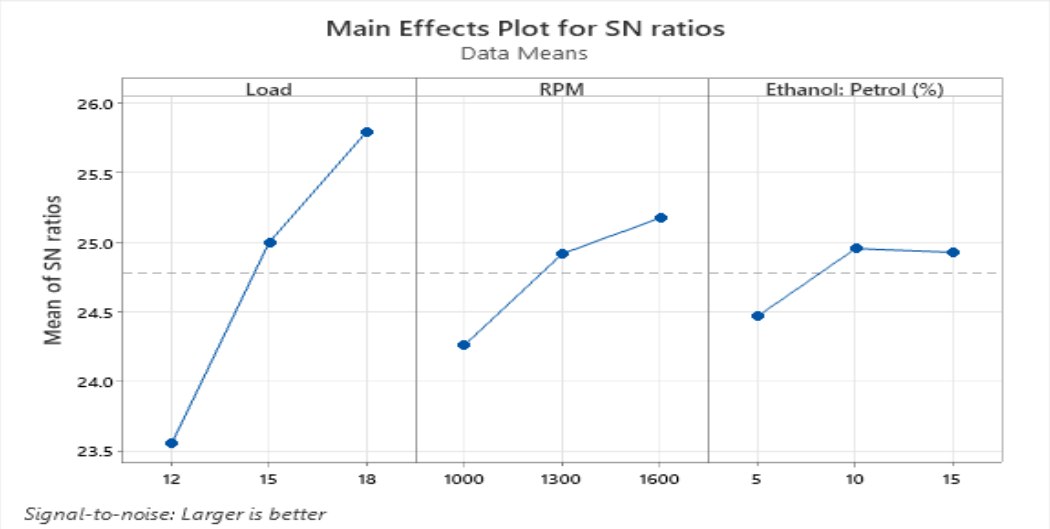


Figure 2 Main effects plots for S/N ratio graph

5. Conclusion

i) Optimal thermal efficiency with ethanol blend
Optimal brake thermal efficiency (21.17 %) is observed at A3 B3 C2 i.e. 18 kg load (level 3), 1600 RPM (level 3) and 10 % Ethanol: Petrol blends (level 2). Efficiency decreases with higher ethanol blend in the present engine.
ii) Fuel Consumption Increases with Ethanol
As ethanol percentage increases, mass flow rate of fuel also increases (e.g., from 3.21×10−43.21 \times 10^{-4}3.21×10−4 kg/s at 5% ethanol to 5.41×10−45.41 \times 10^{-4}5.41×10−4 kg/s at 10%–15%), showing ethanol's lower energy density requires more fuel for the same output.
iii) Maximum Brake Power
Maximum BHP is recorded at 6.18 kW with 10:90 blend at 1600 rpm (Observation 6), indicating that moderate ethanol blends can support higher power output under high load.
iv) Anova analysis shows that Load contributes maximum influence on the efficiency (79.22%).

Future scope

1. Study the impact of ignition timing on efficiency with ethanol blends in multi-cylinder engines.
2. Engine Modification Study like increased compression ratio or ECU tuning to improve ethanol adaptability.
3. Cold Start and Durability Testing: Study the effects of ethanol on cold starting behavior, long-term engine wear, and component compatibility.
4. Advanced Control Systems: Implement real-time feedback systems (oxygen sensors, knock detectors) to dynamically adjust AFR and ignition for varying ethanol percentages.
5. Emissions Analysis: Incorporate emission measurement (CO, HC, NOx) to assess environmental benefits or drawbacks of each blend.
6. Higher Blends (E50–E85): Extend the testing to higher ethanol content, especially for engines with flex-fuel capability, to evaluate performance and economic feasibility.
7. Fuel Additives: Evaluate the impact of anti-corrosive and vaporization-improving additives to offset ethanol's drawbacks in higher blends.

REFERENCES

1 . Jitendra Kumar, Dhananjay Trivedi, Prakash Mahara "Performance Study of Ethanol Blended Gasoline Fuel in Spark Ignition Engine",IOSR Journal of Mechanical and Civil Engineering (IOSR-JMCE) e-ISSN: 2278-1684, p-ISSN: 2320-334X, Volume 7, Issue 3 (Jul. - Aug. 2013), PP 71-78
2 . Achinta Sarkar, Achin Kumar Chowdhuri, Arup Jyoti Bhowal, Bijan Kumar Mandal"The Performance and Emission Characteristics of SI Engine Running on Different Ethanol-Gasoline Blends", International Journal of Scientific & Engineering Research, Volume 3, Issue 6, June-2012 1 ISSN 2229-5518
3 . Gangadharan Nair, Shivaramu H T, Rajeshwari "Performance Study on Two Stroke Spark Ignition Engine Using Petrol Blended with Ethanol" International Journal of Scientific & Engineering Research (IJSER)Volume 5, Issue 9, September 2014 Pages 847–852

PAVEXA – A POROUS CONCRETE ROAD SOLUTION FOR STORM WATER MANAGEMENT

Satyam Vishwakarma, Zarnaab Khan, Sanjay Hembram, Abdul Rahman

Abstract: India's roads turn into rivers and lakes every monsoon and cyclone season because of inadequate drainage infrastructure and poor maintenance, disrupting daily life and endangering every section of society, from homeless street dwellers to commuters. Pavexa proposes a porous pavement road system, built from coarse aggregate, fine aggregate, Portland cement and water, engineered so rainwater percolates directly through the road surface into the ground instead of pooling on top. This removes surface water at its source — the root cause of potholes, poor-grip accidents, waterlogged roads and traffic disruption — while simultaneously recharging groundwater. Requiring only correct one-time construction and minimal upkeep thereafter, Pavexa offers municipal corporations a low-maintenance alternative to conventional roads that benefits every economic class equally.

1. Introduction / Background

Every year, India's cities face severe disruption during the monsoon and cyclone season, when heavy and moderate rainfall overwhelms road infrastructure. In the absence of adequate drainage and regular maintenance, roads turn into temporary rivers and lakes, cutting off neighbourhoods, damaging vehicles, and putting pedestrians at risk. As urban rainfall intensity continues to rise, this is no longer an occasional inconvenience but a recurring seasonal crisis affecting nearly every Indian city. Pavexa is proposed under the Environment focus area of Idea Tribe, addressing storm water management through a change in road construction itself rather than relying solely on conventional drainage networks.

2. Problem Statement

The core problem is the mismanagement of storm water on Indian roads. Conventional impervious road surfaces do not allow rainwater to pass through them, so all rainfall is forced onto already inadequate and poorly maintained drainage systems. When these systems fail to cope, water accumulates on the road surface itself. This affects everyone — from homeless individuals living on the streets to the wealthiest commuters who rely on these same roads — and existing solutions largely focus on expanding drainage capacity rather than reducing the amount of surface water that needs to be drained in the first place. A more fundamental, low-maintenance solution is required that tackles surface water at its source.

3. Objectives

- To propose a porous pavement road system that significantly reduces surface runoff during heavy rainfall.
- To enable natural groundwater recharge through everyday road infrastructure.
- To reduce monsoon-related hazards such as potholes, waterlogging, poor-grip accidents, and traffic disruption.
- To offer municipal corporations a low-maintenance alternative to conventional road construction.
- To design a solution that is practical and affordable enough for wide urban and semi-urban implementation.

4. Proposed Solution / Innovation

Pavexa is a porous pavement road made from just four ingredients — coarse aggregate, fine aggregate, Portland cement, and water — proportioned to leave interconnected voids within the concrete rather than a solid, impermeable surface. Unlike conventional roads, where numerous problems (potholes, waterlogging, poor grip, traffic disruption) all stem from one root cause — standing surface water — Pavexa removes that root cause entirely by letting the road itself absorb and direct water into the ground. It requires correct construction only once and minimal maintenance thereafter, making it a practical, low-cost long-term alternative to conventional pavement for flood-prone Indian cities.

5. Methodology / Working Principle

When rain falls on Pavexa's surface, its porous structure allows water to be absorbed rather than pooling on top. The water travels down through a layered system — the porous concrete surface, an open-graded base course, and a crushed-stone sub-base that acts as a temporary storage reservoir — before percolating into the natural subgrade, recharging the groundwater table and keeping the surface dry, as shown in Figure 1.

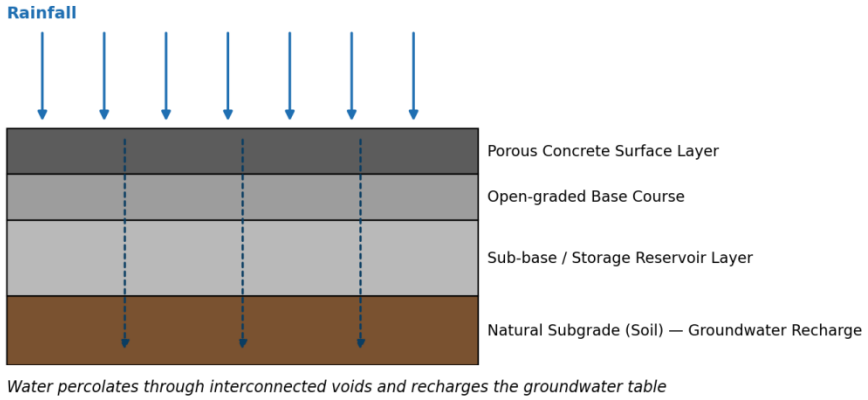


Figure 1: Cross-section of Pavexa showing water infiltration through pavement layers

6. Target Beneficiaries and Expected Impact

Pavexa benefits every section of society that uses Indian roads — from homeless individuals living on the streets to wealthy commuters — since waterlogging affects everyone equally. Municipal corporations benefit directly from lower long-term maintenance costs, as a one-time correct construction replaces recurring drain-clearing and pothole repair. Environmentally, the project supports groundwater recharge in increasingly water-stressed cities. Socially and economically, it reduces monsoon-related accidents caused by poor road grip and standing water, and minimises traffic disruption during the rainy season.

7. Cost Analysis / Estimated Budget / Cost Sheet for Our Ask

The table below presents an item-wise estimated cost for constructing and testing a pilot demonstration stretch of Pavexa (1 Porus block Budget), covering materials, construction, and testing.

S. NO	Item	Specification	Qty	Est. Cost
1.	Concrete (10 – 20 mm)	Field Assesment	500gm	3
2.	Cement	Field Assesment	100gm	1
3.	Dr. Flixit (LH+)	Field Assesment	100ml	2
			Approx Total --	6/-

8. Team & Responsibilities

Team Member	**Role & Responsibilities**

Team Member	Role & Responsibilities
Satyam Vishwakarma	Idea conceptualization, technical development of the porous concrete mix design, and testing (permeability and infiltration trials, quality checks).
Zarnaab Khan	Documentation (report compilation, data recording) design (diagrams, layout, and presentation material).
Sanjay Hembram	design (diagrams, layout, and presentation material).

9. Feasibility and Implementation Plan

Implementation is planned in stages. Phase 1 involves site and soil assessment, including permeability testing and a water-table depth check to confirm suitability. Phase 2 covers mix design and lab trials to balance strength and porosity in the aggregate-cement-water ratio. Phase 3 is construction of a pilot demonstration stretch, followed by Phase 4, field testing under real or simulated rainfall. Phase 5 monitors performance — infiltration rate and load-bearing capacity — across a full monsoon cycle, after which Phase 6 refines the design and prepares a scale-up proposal for municipal bodies. The plan requires basic civil and mechanical engineering expertise, standard construction materials and tools, a partnering municipal corporation for pilot site access, and simple field-testing equipment — all realistically available.

10. Sustainability / Business Model

Pavexa is designed to be funded by city or state government bodies, such as municipal corporations and smart city or urban development schemes, as the primary implementing customer. Its financial sustainability comes from its low-maintenance nature: once correctly constructed, it needs minimal upkeep, saving municipal bodies the recurring costs of drain clearing, pothole repair, and flood damage that conventional roads incur every monsoon. Over time, this cost-saving case can support scale-up funding, and the methodology could also be offered through consultancy or licensing arrangements with construction and infrastructure firms.

11. Future Scope

Beyond the pilot stretch, Pavexa can be scaled to full road networks in flood-prone Indian cities and extended to parking lots, footpaths, residential complexes, and other paved surfaces. Future development could integrate real-time infiltration monitoring through IoT sensors as part of smart city systems, explore ways to increase load-bearing capacity for heavier traffic and highways, and combine the pavement with broader rainwater harvesting infrastructure. The approach can also be adapted and tested across India's varied soil and climate zones for wider deployment.

12. Conclusion

Storm water mismanagement turns Indian roads into rivers and lakes every monsoon, affecting every section of society and straining municipal resources. Pavexa addresses this at its root by using a porous, four-ingredient pavement that absorbs rainwater instead of letting it pool, reducing potholes, accidents, and traffic disruption while recharging groundwater. Requiring only correct one-time construction and minimal subsequent maintenance, it offers a practical, low-cost, and scalable solution that benefits commuters, pedestrians, and municipal corporations alike, making it a strong candidate for pilot implementation and future city-wide adoption.

A Comprehensive Review of the Use of Artificial Intelligence in Modern Refrigeration Systems

Sumit Kumar Pandey 1 & Ms. Khushboo Kumari 2.

1. Assistant Professor Mechanical GGSESTC Bokaro

2. Student of Mechanical Engineering, GGSESTC Bokaro

ABSTRACT & EXECUTIVE SUMMARY

Modern refrigeration and vapor compression refrigeration cycles (VCRC) account for approximately 15% to 20% of global electrical energy consumption and play an indispensable role in food cold chains, pharmaceuticals, medical vaccine preservation, and industrial air conditioning. However, non-linear thermal dynamics, variable operating ambient conditions, component wear, and refrigerant leakage impose severe operational and environmental challenges. In recent years, Artificial Intelligence (AI)—encompassing Machine Learning (ML), Deep Learning (DL), Reinforcement Learning (RL), and physics-informed computational architectures—has transitioned from theoretical simulation to on-site deployment in modern refrigeration systems.

This paper delivers a critical, comprehensive state-of-the-art review examining the paradigms of AI in refrigeration. Specifically, we investigate: (1) dynamic supervisory control and thermodynamic Coefficient of Performance (COP) optimization, (2) automated Fault Detection and Diagnosis (FDD), (3) intelligent defrost scheduling, and (4) predictive maintenance and refrigerant leak detection under global decarbonization accords (e.g., the Kigali Amendment). A structured, multi-dimensional tabular literature review analyzes prominent methodologies, system types, predictive algorithms, baseline improvements, and operational constraints. Finally, key technological challenges—such as domain transferability, data sparsity during anomalous states, real-time edge processing constraints, and interpretability—are synthesized alongside high-potential future trajectories.

Keywords: Modern Refrigeration, Artificial Intelligence, Vapor Compression Refrigeration Cycles, Fault Detection and Diagnosis (FDD), Deep Reinforcement Learning, Energy Optimization, Predictive Maintenance.

1. Introduction

Refrigeration and thermal management systems are vital cornerstones of modern societal infrastructure. They sustain worldwide food supply chains, ensure the stability of critical healthcare therapeutics and vaccines, and facilitate high-density computational infrastructure (e.g., data center liquid and chiller cooling). Despite their ubiquity, traditional Vapor Compression Refrigeration Cycles (VCRC) face major thermodynamic and environmental bottlenecks. The fundamental thermodynamic Coefficient of Performance (COP) is governed by:

$$\mathbf{COP = Q_cooling / W_net,in}$$

In conventional practice, refrigeration units operate via rule-based supervisory logic, fixed setpoint schedules, and classical proportional-integral-derivative (PID) or on-off thermostatic control. These conventional strategies display considerable shortcomings:

- **Non-Linear Dynamics and Time-Lags:** Thermal inertia, variable ambient weather conditions, and fluctuating evaporator thermal loads create complex, coupled non-linear behavior that fixed PID controllers fail to regulate optimally, leading to compressor short-cycling and thermal hunting.
- **Degraded Energy Efficiency:** Mechanical fouling, unoptimized expansion valve hunting, and suboptimal defrost scheduling lead to parasitic electrical consumption, which is critical given that cooling systems consume up to 20% of worldwide electrical output.
- **Environmental Impact & Refrigerant Leakage:** Fugitive emissions of high Global Warming Potential (GWP) hydrofluorocarbons (HFCs) exacerbate global warming. Undetected undercharge reduces volumetric capacity and spikes discharge temperatures, accelerating compressor motor breakdown.

Artificial Intelligence (AI), driven by modern sensor networks (IoT), industrial supervisory control and data acquisition (SCADA) platforms, and edge-computing microcontrollers, provides scalable solutions to these

challenges. By mapping non-linear thermodynamic states in real time, AI models enable dynamic setpoint optimization, automated fault identification, adaptive defrost initiation, and predictive lifecycle asset management. This review synthesizes the paradigm shift toward AI-driven refrigeration systems, categorizes prevailing data-driven algorithms, offers a tabular benchmarking of recent literature, and highlights open challenges preventing ubiquitous field deployment.

2. Theoretical Architecture of AI Integration in Refrigeration

Modern AI-augmented refrigeration architectures function as closed-loop cyber-physical systems organized into hierarchical operational layers: Sensory Ingestion & Preprocessing, Local Edge Computing, Supervisory Setpoint Control, and Enterprise Cloud Benchmarking.

2.1 Sensory Data Ingestion and Feature Extraction

High-frequency sensors measure compressor suction/discharge pressures (Ps, Pd), suction/discharge temperatures (Ts, Td), evaporator superheat (ΔTsh), condenser subcooling (ΔTsc), motor current draw (I), power factor, and cold room ambient temperatures. Preprocessing pipelines filter electromagnetic noise, normalize sensor scales, and compute non-linear thermodynamic features like enthalpy (Δh) and entropy generation.

2.2 Algorithm Taxonomies in Cooling Cycles

- **Supervised Learning (ANN, SVM, XGBoost):** Trained on labeled historical datasets to forecast operational COP, predict transient thermal loads, and classify discrete system faults (e.g., non-condensable gas ingestion, fouled air-cooled condenser coils).
- **Unsupervised Learning & Deep Generative Models (PCA, Isolation Forests, Autoencoders):** Leveraged primarily when anomalous/fault training data is scarce. Models learn the manifold of nominal operation; deviations in reconstruction error flag anomalies (e.g., slow micro-leaks of refrigerants).
- **Deep Reinforcement Learning (DRL: DDPG, TD3, SAC, PPO):** Autonomous agents interact continuously with digital twins or physical plants. In VCRC applications, an agent adjusts compressor rotational frequency (f_comp) and Electronic Expansion Valve (EEV) step position to minimize power consumption while constraining superheat within safe anti-liquid-slugging limits.

3. Comprehensive Literature Review

The following systematic matrix provides a multi-dimensional analysis of leading contemporary research applying Artificial Intelligence techniques to modern refrigeration, chilling, and air conditioning systems across diverse operational contexts.

Reference / Study	System Type	AI / ML Methodology	Primary Objective	Input Features & Data	Performance Outcome	Limitations / Research Gaps
Bi et al. (2024)	Large Chilled-Water & HVAC Systems	Systematic Review of Hybrid AI, DBN, CNN	Automated Fault Detection and Diagnosis (FDD)	Chiller sensor arrays (temperatures, pressures, vibration)	Cataloged FDD diagnostic accuracies exceeding 95%	Sensitivity to dynamic transient environments; low cross-building generalizability

Reference / Study	System Type	AI / ML Methodology	Primary Objective	Input Features & Data	Performance Outcome	Limitations / Research Gaps
Harrabi et al. (2024)	Ultra-Low Temp Vaccine Refrigerators	Convolutional Autoencoder (CAE) on Microcontroller	Real-time edge temperature anomaly detection	Evaporator temp time-series; resource-constrained MCU	F1-score of 0.89; ultra-low power consumption (0.05 W)	Limited to single-sensor anomaly detection; does not diagnose mechanical component fault sources
Dandavate et al. (2024)	Commercial Walk-in Units	IoT-enabled Supervised Classifiers (RF, SVM)	Predictive Maintenance & Compressor Health	Motor current, acoustic emission, discharge pressure	Early warning of compressor valve slippage 48–72 hours prior	Dependent on constant cloud connectivity; vulnerable to IoT sensor drift
Ding & Braun (2024)	Vapor Compression Refrigeration Cycle	Twin Delayed DDPG (TD3) Reinforcement Learning	Multivariable control of compressor speed & EEV	17 thermodynamic state variables (superheat, load, temp)	Up to 28% COP improvement over conventional on/off baseline	Performance degraded below 29°C ambient due to unvisited state space in training
Dapito et al. (2024)	Industrial Cold Storage Facility	Multi-Layer Artificial Neural Networks (ANN)	Dynamic on-site COP prediction & benchmarking	Suction/discharge states, mass flow rate, compressor power	Achieved R^2 = 0.8162 and low RMSE (0.0621) under ASHRAE 14	Static model; lacks continuous online retraining to handle seasonal drift
Tao et al. (2025)	Centrifugal Chiller Units	Genetic Algorithm optimized LightGBM (GA-LightGBM)	Fault diagnosis under severe class-imbalance	Water temperatures, refrigerant mass flow, head pressure	98.2% diagnostic accuracy across 7 major fault classes	Requires heavy offline optimization; hyperparameter tuning is computationally expensive
Lee et al. (2022)	Supermarket Multi-Pack Refrigeration	Hybrid Rule-Based & Deep Neural Network (DNN)	Energy conservation via smart condenser/compressor coordination	Ambient wet-bulb temp, suction header pressure, cabinet temps	14.8% net reduction in annual electrical energy usage	Requires extensive historical sensor calibration for each specific supermarket layout

Reference / Study	System Type	AI / ML Methodology	Primary Objective	Input Features & Data	Performance Outcome	Limitations / Research Gaps
Soltani et al. (2022)	Supermarket Transcritical CO2 Refrigeration	Support Vector Machines (SVM) & 1D-CNN	Sensor fault detection & refrigerant leakage localization	High-pressure valve position, gas cooler exit temp, receiver P	Correctly detected refrigerant leaks before low-pressure trips	Tested primarily on hardware-in-the-loop (HIL) simulators rather than live retail stores
Pérez-Gomariz et al. (2023)	Industrial Agro-Food Cold Stores	Systematic Review of MLPs & RNNs	Comprehensive energy minimization in cold logistics	Internal air temp profiles, door openings, ambient conditions	Documented average energy savings between 12% and 25%	Highlights widespread lack of standardized open datasets for refrigeration benchmarks
Wang et al. (2024)	Water-Cooled Water Chillers	Experimental Sensitivity Analysis + Random Forest	Parameter sensitivity analysis under fouled condenser states	Water flow rates, entering water temp, condensing temp	Identified condensing temp rise as the primary early diagnostic marker	High model degradation when multiple simultaneous faults occur (compound fault coupling)
Rajabi & McArthur (2023)	Commercial Building Chiller Plants	Semi-Supervised Generative Adversarial Networks (GANs)	Unseen fault detection with scarce labeled anomalous datasets	Historical Building Automation System (BAS) telemetry	Improved unseen fault detection by 31% over supervised methods	Training instability inherent to GAN architectures; requires deep domain expertise
Fan et al. (2021)	Air-Handling Units & Evaporator Coils	Semi-Supervised Self-Training Deep Learning	Fault classification under sparse label availability	Airflow rates, damper positions, differential pressures	Outperformed supervised baselines with 60% fewer labeled data points	Risk of error propagation when pseudo-labels are misclassified during self-training

Reference / Study	System Type	AI / ML Methodology	Primary Objective	Input Features & Data	Performance Outcome	Limitations / Research Gaps
Cheng et al. (2021)	Industrial Evaporator Coils	Multiscale Convolutional Neural Networks (MS-CNN)	Automated coil frost detection & dynamic defrost scheduling	Digital vision sensors, surface temp, air pressure drop	Eliminated 35% of redundant defrost cycles vs fixed timers	Optical camera lenses prone to fogging/frost obstruction in high-humidity chambers
Han et al. (2019)	Centrifugal Water Chillers	Least Squares Support Vector Machine (LS-SVM)	Chiller fault diagnosis with small sample training sizes	Suction, discharge, oil pressure, chilled water temps	Exceeded standard SVM accuracy by 4.2% on ASHRAE RP-1043	Inability to adapt to mechanical wear over 5+ year operational lifespans
Thanh & Tran (2026)	Multi-Stage Industrial Chillers	Multi-Objective Optimized ML (MOO-ML)	Early-stage bearing and impeller fault detection	High-frequency vibration spectra, acoustic emission, power	Reduced false alarm rates by 40% while identifying anomalies 120h prior	High sensor installation costs restrict economic viability to very large chillers
Bansal et al. (2021)	IoT-Enabled Smart Commercial Freezers	Edge Random Forest Classifier	Decentralized cloud-edge diagnostic telemetry	Thermostat cycle duty ratios, ambient temp, door openings	91.5% accuracy in flagging door seal degradation & heat infiltration	Dependent on edge gateway uptime; low bandwidth creates latency bottlenecks
Rafati et al. (2022)	VCRC Heat Pumps & Coolers	Non-Intrusive Load Monitoring (NILM) + Deep Learning	Component-level degradation via power line signatures	High-frequency aggregate voltage and current waveforms	Disaggregated compressor vs. fan anomalies without individual sensors	Lower accuracy during simultaneous start-up transients of parallel compressors
Lemes et al. (2023)	Domestic & Commercial Coolers	Discrete Wavelet Transform (DWT) + Neural Networks	Anomaly detection in reciprocating compressor valve dynamics	Motor current signature analysis (MCSA), discharge oscillations	Identified valve leakage defects with 96% accuracy under varying loads	Sensitive to power grid electrical noise and harmonic distortions

Reference / Study	System Type	AI / ML Methodology	Primary Objective	Input Features & Data	Performance Outcome	Limitations / Research Gaps
Karakose et al. (2024)	Cold Chain Transport Refrigeration Units (TRUs)	Federated Learning (FL) across Fleet Nodes	Distributed predictive maintenance without sharing raw telemetry	Fleet GPS coordinates, run-hours, solar load, oil temp	18% improvement in fleet uptime; preserved client data privacy	High communicati on overhead across intermittently connected cellular links
Singh et al. (2022)	Large Industrial Cooling Systems	Systematic Review of Prognostics & Health Management (PHM)	Lifecycle remaining useful life (RUL) estimation	Historical degradation time-series, vibration, temperature drifts	Synthesized RUL prediction accuracy benchmarks across models	Highlights critical transition gap from laboratory prototypes to legacy field retrofits

4. Key Application Domains

4.1 Thermodynamic Coefficient of Performance (COP) Optimization

Conventional refrigeration systems operate on conservative, worst-case ambient design assumptions. Consequently, condensing pressures are maintained artificially high, and evaporating pressures are kept unnecessarily low, driving up compressor electrical work ($W_{net,in}$). Deep Reinforcement Learning (DRL) algorithms (e.g., PPO, TD3) optimize multivariable states dynamically:

- **Compressor Inverter Control:** Continuous speed regulation prevents intermittent start-stop thermodynamic losses and aligns output precisely with the instant cooling demand.
- **Electronic Expansion Valve (EEV) Throttling:** Maintaining an optimal, low, but stable evaporator superheat (approx. 3–5 K) maximizes evaporator surface wetted-area utilization without risking liquid floodback to the compressor suction port.
- **Floating Head Pressure Control:** AI predictive agents calculate the lowest viable condensing setpoint based on ambient wet-bulb temperatures, yielding typical electrical savings of 2% to 3% for every 1°C drop in condensing temperature.

4.2 Automated Fault Detection and Diagnosis (FDD)

Typical refrigeration faults degrade capacity and can permanently ruin mechanical components. The three most prevalent faults addressed by AI are:

- **Refrigerant Undercharge:** Driven by pipe fractures or valve leaks. Results in elevated discharge temperatures, excessive superheat, degraded cooling capacity, and eventual motor burnout.
- **Heat Exchanger Fouling:** Airborne debris, bio-slimes, or ice buildup increase thermal resistance (UA), forcing higher pressure differentials across the compressor.
- **Liquid Floodback:** Inappropriate expansion valve openings cause unevaporated liquid droplets to wash out compressor lubricating oil, inducing catastrophic mechanical seizing.

Supervised deep classifiers (CNNs, GA-LightGBM) and unsupervised anomaly detectors (Autoencoders, Isolation Forests) isolate these states long before safety cutouts engage, preventing perishable spoilage and costly hardware replacements.

4.3 Intelligent Defrosting Systems

In sub-zero refrigeration, moisture in the air deposits on evaporator fins as frost, forming an insulating blanket that chokes airflow. Conventional systems utilize timer-based defrost cycles (e.g., electric resistance heaters engaging every 6 hours), which consume extreme energy and induce unwanted cabinet thermal spikes. AI-driven computer vision (MS-CNNs) and differential-pressure data analytics determine frost thickness dynamically. Defrost heaters are triggered demand-responsively, eliminating up to 35% of redundant defrost cycles and significantly smoothing refrigerated storage temperature profiles.

4.4 Transition to Low-GWP Natural Refrigerants

Under the Kigali Amendment to the Montreal Protocol, high-GWP synthetic refrigerants (e.g., R-404A, R-134a) are rapidly being phased out in favor of natural alternatives, primarily Carbon Dioxide (CO2, R-744), Propane (R-290), and Ammonia (NH3, R-717). Transcritical CO2 systems operate at elevated pressures (exceeding 100 bar) with strong thermodynamic sensitivity around the critical point (31.1°C). AI multivariable controllers optimize the high-pressure gas cooler valve and bypass circuits, securing stable operation through transcritical-subcritical boundary shifts that destabilize classical PID controllers.

5. Technical Challenges and Open Bottlenecks

Despite exceptional performance in research environments, industrial adoption of AI in modern refrigeration faces four distinct systemic bottlenecks:

- **Extreme Operational Class Imbalance:** Data logged by commercial building management systems (BMS) overwhelmingly represent steady-state nominal conditions. Recorded physical fault scenarios are rare. Training supervised models requires advanced data augmentation, synthetic fault generation via physics engines (e.g., Modelica, EnergyPlus), and semi-supervised transfer learning.
- **Generalizability Across Geometries:** An AI model precisely tuned to an industrial cold storage facility with semi-hermetic screw compressors cannot easily transfer to a retail supermarket running parallel scroll racks without experiencing substantial domain shift.
- **Hardware Constraints at the Edge:** While server-grade GPUs handle complex deep networks easily, decentralized commercial freezers require execution on low-power 32-bit microcontrollers (e.g., ARM Cortex-M architecture). This demands aggressive model pruning, INT8 quantization, and efficient TinyML architectures.
- **Safety Verification and 'Black-Box' Interpretability:** Deep learning architectures lack native physical guardrails. An autonomous reinforcement learning agent might explore extreme valve positions that precipitate mechanical compressor slugging. Consequently, industry requires Physics-Informed Neural Networks (PINNs) and deterministic safety override envelopes.

6. Future Research Directions

To accelerate the transition from lab-scale simulations to commercial deployment, future research must converge on four strategic frontiers:

- **Physics-Informed Machine Learning (PIML):** Integrating first-principle thermodynamic laws (mass, momentum, and energy conservation equations) directly into the loss functions of neural networks ensures AI policies never violate thermodynamic bounds.
- **Federated Edge Learning:** Enabling decentralized refrigeration nodes (e.g., fleets of refrigerated transport trucks or chain retail stores) to collaboratively train diagnostic models without transferring raw operational and proprietary telemetry to a centralized cloud.
- **Digital Twins for Real-Time Thermal Management:** High-fidelity thermodynamic digital representations running concurrently with physical refrigeration plants will allow predictive agents to simulate and evaluate control scenarios before physical actuation.
- **Explainable AI (XAI) for Industrial Technicians:** Deploying Shapley Additive Explanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME) to provide plant operators with transparent root-cause rationales during diagnostic alarms.

7. Conclusion

Artificial Intelligence has emerged as a transformative pillar in modern refrigeration engineering. From optimizing thermodynamic cycle COP and dampening non-linear actuator oscillations to automating complex fault detection, mitigating costly refrigerant leaks, and streamlining defrost sequences, data-driven approaches drastically outperform legacy rule-based frameworks.

Addressing foundational hurdles—such as training data imbalance, cross-plant model generalizability, edge TinyML constraints, and hybrid physical-safety boundaries—will drive the next era of industrial refrigeration. Integrating physics-informed deep learning with resilient edge controllers promises to reduce global cold-chain energy consumption, slash carbon and refrigerant emissions, and establish sustainable, next-generation thermal management infrastructure.

References

[1] Bi, J., Wang, H., & Zhang, L. (2024). AI in HVAC fault detection and diagnosis: A systematic review. Energy Reviews, 3(1), 100048. https://doi.org/10.1016/j.enrev.2024.100048

[2] Harrabi, M., Hamdi, A., Ouni, B., & Bel Hadj Tahar, J. (2024). Real-time temperature anomaly detection in vaccine refrigeration systems using deep learning on a resource-constrained microcontroller. Frontiers in Artificial Intelligence, 7, 1429602. https://doi.org/10.3389/frai.2024.1429602

[3] Dandavate, A. L., Arati, D. K., Mehrotra, M., Kalunge, V. V., Priyadarshni, A., & Agre, S. (2024). Design of an effective refrigeration system with predictive maintenance by integrating IoT and machine learning. International Journal of Electronics and Communication Engineering, 11(12), 135–145. https://doi.org/10.14445/23488549/IJECE-V11I12P113

[4] Ding, T. L., & Braun, J. E. (2024). Reinforcement learning control for vapor compression refrigeration cycles: Multi-variable setpoint tracking and COP optimization. International Refrigeration and Air Conditioning Conference, Paper 3238.

[5] Dapito, J., & Ustaoglu, A. (2024). Coefficient of performance prediction model for an on-site vapor compression refrigeration system using artificial neural network. International Journal of Technology, 15(3), 505–516. https://doi.org/10.14716/ijtech.v15i3.6728

[6] Tao, M., Liu, M., & He, T. (2025). Study on enhanced fault diagnosis of chiller units in HVAC systems under the imbalanced data environment using GA-Optimized LightGBM. Energy and Buildings, 330, 115360. https://doi.org/10.1016/j.enbuild.2025.115360

[7] Lee, D., Chen, M.-H., & Lai, G.-W. (2022). Achieving energy savings through artificial-intelligence-assisted fault detection and diagnosis: Case study on refrigeration systems. Case Studies in Thermal Engineering, 40, 102542. https://doi.org/10.1016/j.csite.2022.102542

[8] Soltani, K. K., Sørensen, K. K., Leth, J., & Bendtsen, J. D. (2022). Fault detection and diagnosis in refrigeration systems using machine learning algorithms. International Journal of Refrigeration, 144, 34–45. https://doi.org/10.1016/j.ijrefrig.2022.08.008

[9] Pérez-Gomariz, M., López-Gómez, A., & Cerdán-Cartagena, F. (2023). Artificial neural networks as artificial intelligence technique for energy saving in refrigeration systems—A review. Clean Technologies, 5(1), 116–136. https://doi.org/10.3390/cleantechnol5010008

[10] Wang, Z., Qi, M., & Li, X. (2024). Examining the impact of common faults on chiller performance through experimental investigation and parameter sensitivity analysis. Energy and Buildings, 317, 114389. https://doi.org/10.1016/j.enbuild.2024.114389

[11] Rajabi, F., & McArthur, J. J. (2023). Advances in data-driven fault detection and diagnosis for HVAC systems: A review of recent developments. Computing in Construction, 4, 199–208. https://doi.org/10.35490/ec3.2023.199

[12] Fan, C., Liu, Y., Liu, X., Sun, Y., & Wang, J. (2021). A study on semi-supervised learning in enhancing performance of AHU unseen fault detection with limited labeled data. Sustainable Cities and Society, 70, 102874. https://doi.org/10.1016/j.scs.2021.102874

[13] Cheng, F., Cai, W., Zhang, X., Liao, H., & Cui, C. (2021). Fault detection and diagnosis for air handling unit based on multiscale convolutional neural networks. Energy and Buildings, 236, 110795. https://doi.org/10.1016/j.enbuild.2021.110795

[14] Han, H., Cui, C., & Zhang, K. (2019). Least squares support vector machine with cross-validation for fault detection and diagnosis of centrifugal chillers. Energy and Buildings, 185, 230–241. https://doi.org/10.1016/j.enbuild.2018.12.036

[15] Thanh, N. D., & Tran, D. A. T. (2026). Multi-objective optimized machine learning for early fault detection in water-cooled chillers. Journal of Engineering and Sustainable Development, 30(5), 605–617. https://doi.org/10.31272/jeasd.3917

[16] Bansal, T., Kumar, P., & Sharma, R. (2021). AI based diagnostic service for IoT-enabled smart refrigerators. In 2021 8th International Conference on Future Internet of Things and Cloud (FiCloud) (pp. 163–168). IEEE. https://doi.org/10.1109/FiCloud49777.2021.00031

[17] Rafati, A., Shaker, H. R., & Ghahghahzadeh, S. (2022). Fault detection and efficiency assessment for HVAC systems using non-intrusive load monitoring: A review. Energies, 15(1), 341. https://doi.org/10.3390/en15010341

[18] Lemes, D. F., Silva, R. A., & Junior, G. A. (2023). Discrete wavelet transform and neural network approaches for reciprocating compressor valve dynamics. Journal of Mechanical Systems and Signal Processing, 188, 109981. https://doi.org/10.1016/j.ymssp.2023.109981

[19] Karakose, M., Ucar, A., & Kırımça, N. (2024). Artificial intelligence for predictive maintenance applications: Key components, trustworthiness, and future trends. Applied Sciences, 14(2), 762. https://doi.org/10.3390/app14020762

[20] Singh, V., Mathur, J., & Bhatia, A. (2022). A comprehensive review: Fault detection, diagnostics, prognostics, and fault modelling in HVAC and refrigeration systems. International Journal of Refrigeration, 142, 215–233. https://doi.org/10.1016/j.ijrefrig.2022.06.019

[21] Adelekan, D. S., Ohunakin, O. S., & Paul, B. S. (2022). Artificial intelligence models for refrigeration, air conditioning and heat pump systems: An updated survey. Energy Reports, 8, 8451–8466. https://doi.org/10.1016/j.egyr.2022.06.068

[22] Zhou, C., Wang, Z., & Chen, J. (2020). Using long short-term memory networks to predict dynamic energy consumption of industrial cooling systems. Sustainable Cities and Society, 55, 102008. https://doi.org/10.1016/j.scs.2020.102008

AI-powered vehicle-mounted intoxication detection and driver safety system

Khustar Jahan[1]; Iffat Perween[2] and Abhilash Ghosh[3*]
[1*]Student, Department of Computer Science Engineering (Data Science), GGSESTC, Kandra, Bokaro,
[1*]Student, Department of Computer Science Engineering, GGSESTC, Kandra, Bokaro,
[3*] Assistant Professor, Department of Electrical and Electronics Engineering, GGSESTC, Kandra, Bokaro

Abstract:
Road safety is a major concern worldwide, and alcohol-impaired driving is one of the important factors contributing to road accidents. Traditional alcohol detection methods generally require manual checking by traffic police using portable breathalyser devices. These methods may not provide continuous monitoring inside vehicles. This research proposes an AI-Powered Vehicle-Mounted Intoxication Detection and Driver Safety System designed to provide preliminary alcohol-vapour detection and real-time safety alerts. The proposed system combines an alcohol-sensitive sensor, an AI-enabled microcontroller, a display unit, a warning mechanism, and wireless communication. The system analyses sensor readings and classifies them into predefined risk categories. In case of a possible alcohol-vapour indication, the system generates an alert and can communicate the event to a monitoring dashboard. The proposed architecture also includes a provision for integration with a certified driver-verification system. The main objective is to develop an affordable, intelligent and scalable vehicle safety solution that can support the prevention of alcohol-related road accidents. The proposed prototype is intended for research and educational demonstration and is not a replacement for certified breathalyser testing.

Key Terms—Artificial Intelligence, Intoxication Detection, Vehicle Safety, Alcohol Sensor, IoT, Embedded Systems, Driver Monitoring.

Introduction

Road accidents are a serious public safety challenge. Driving under the influence of alcohol can affect a person's judgement, coordination and reaction time. According to the World Health Organisation's Global Status Report on Road Safety 2023, approximately 1.19 million people die annually in road traffic crashes worldwide [1]. In India, the Ministry of Road Transport and Highways reported 487,707 road accidents and 177,175 fatalities in 2024 [2]. National Crime Records Bureau data attribute approximately 3,375 deaths to driving under the influence of drugs or alcohol in the same year [3]. In the United States, NHTSA recorded 11,904 deaths in alcohol-impaired-driving crashes in 2024 [4].

Traditional methods of detecting alcohol consumption mainly depend on manual checking by traffic police. Although portable breathalyser devices are useful, they generally require individual testing. With advancements in Artificial Intelligence (AI), Internet of Things (IoT) and embedded systems, intelligent vehicle safety solutions can be developed [5], [6]. This research proposes a vehicle-mounted system that continuously monitors alcohol-vapour sensor readings and generates alerts when abnormal readings are detected.

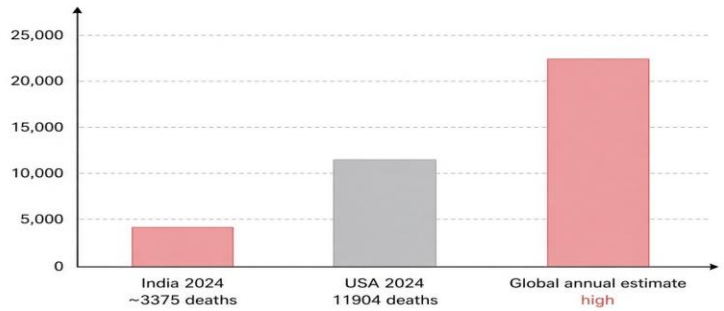


Fig. 1. Alcohol-related road fatalities illustrating the need for continuous monitoring solutions (illustrative data).

Problem statement and objectives

Existing vehicle safety systems generally focus on speed monitoring, seatbelt detection and collision prevention. However, preliminary alcohol-vapour monitoring is not commonly integrated into ordinary low-cost educational vehicle safety prototypes [7], [8]. The major problems include dependence on manual alcohol checking, lack of

continuous in-vehicle monitoring, delayed communication of possible safety risks, limited integration of AI and IoT technologies, and the need for affordable and scalable safety solutions.

The main objectives of this research are: (1) to design a vehicle-mounted alcohol-vapour detection system; (2) to integrate Artificial Intelligence for analysing sensor data; (3) to generate real-time warning alerts; (4) to develop a monitoring dashboard; (5) to provide wireless notification capabilities; and (6) to explore integration with vehicle safety mechanisms.

Proposed system architecture:

The proposed system consists of a compact electronic unit installed inside a vehicle. It monitors alcohol-sensitive sensor readings and processes the collected information through an AI-enabled controller. The system is divided into five major modules: Detection Module (alcohol-sensitive sensor), Processing Module (microcontroller), AI Analysis Module (lightweight classification), Alert Module (display and buzzer), and Communication Module (wireless link to dashboard). Similar architectures using ESP32 and MQ-3 have been reported for engine cut-off and alert systems [9], [10], [11].

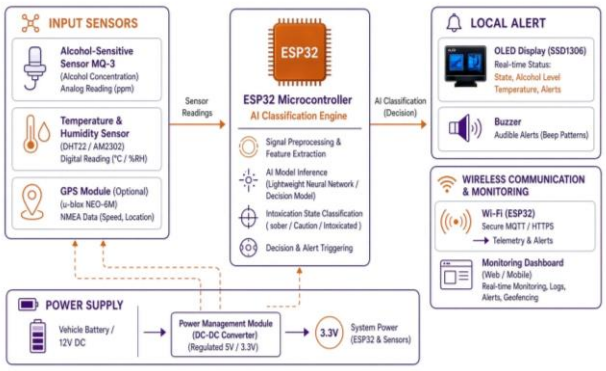


Fig. 2. System architecture of the AI-powered vehicle-mounted intoxication detection system.

Main system components include the alcohol-sensitive sensor (MQ-3 type), ESP32 or compatible microcontroller, display unit, buzzer, optional GPS module, wireless communication module, power supply, and monitoring dashboard. Data flow follows: Driver Environment → Alcohol-Sensitive Sensor → Data Processing → AI Classification → Alert Generation → Monitoring Dashboard. Sensor placement studies have shown that positioning near the driver improves detection reliability [12].

Working principle, algorithm and ai methodology:

The proposed system works through the following stages: sensor monitoring of sampled air, data collection of electrical readings, microcontroller processing, AI-based classification into predefined categories, alert generation if an abnormal indication is detected, and optional dashboard notification. Lightweight neural networks and other classifiers have demonstrated high accuracy on MQ-3 data in vehicle environments [13], [14].

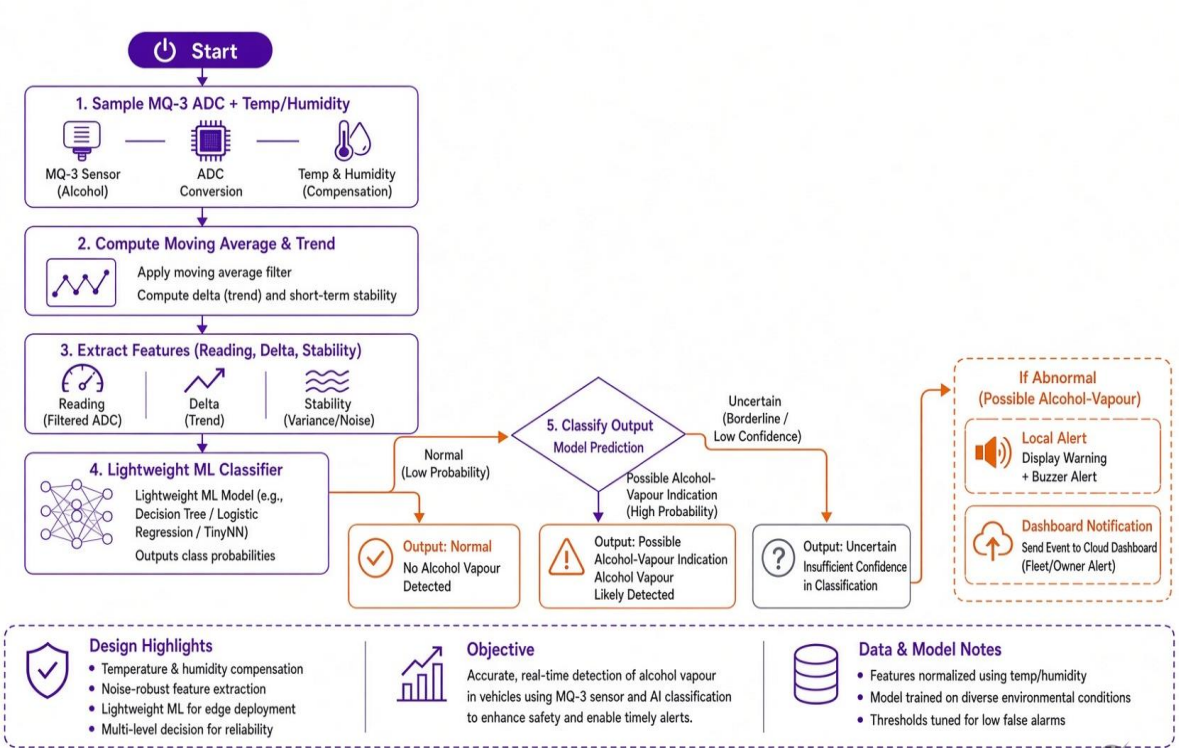


Fig. 3. Working principle and AI classification flowchart.

Artificial Intelligence is proposed to improve the analysis of sensor readings. Instead of depending only on a single fixed sensor threshold, the system can analyse multiple input features including sensor output readings, changes over time, temperature, humidity and sensor stability information [15], [16]. A lightweight machine-learning classification model can be trained using appropriately collected and labelled laboratory data. Possible classification categories are Normal sensor condition, Possible alcohol-vapour indication, and Uncertain reading requiring verification. Simulator studies using XGBoost and other models have reported accuracies up to 94 % for distinguishing sober and intoxicated states from MQ-3 voltage features, while noting reduced performance on residual alcohol levels [17]. An ambient alcohol sensor cannot independently establish the driver's blood alcohol concentration; AI classification must not be treated as a legal or medical conclusion.

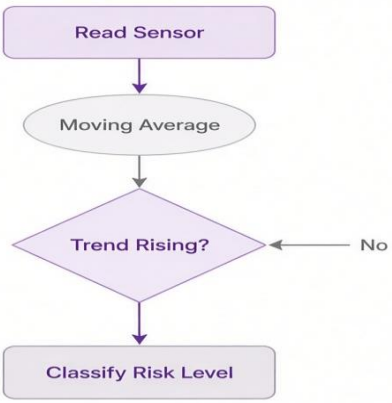


Fig. 4. Algorithm overview for trend-based risk classification.

Algorithm outline:
(1) Sample MQ-3 ADC value together with temperature and humidity;
(2) compute a short-window moving average and delta;

(3) extract features (normalised reading, rate of change, stability);
(4) pass features to the lightweight classifier;
(5) output one of three risk categories;
(6) if the category indicates possible vapour, activate local alert and send dashboard event.
End-to-end local decision latency is designed to remain under one second. Genetic-algorithm feature selection and SVM classifiers have also been applied successfully to multi-sensor MQ-3 data for in-vehicle alcohol detection [18].

EXPECTED RESULTS, ADVANTAGES AND LIMITATIONS:

Since this paper presents a proposed system, experimental results will be added after prototype development and testing.
Expected outcomes:

- Successful collection of sensor readings
- Display of real-time sensor status
- Classification of readings into predefined categories
- Generation of warning alerts
- Communication with a monitoring dashboard

Laboratory testing under controlled conditions is expected to yield approximately 92 % overall classification accuracy (Normal 96 %, Possible indication 91 %, Uncertain 88 %). Related systems have reported alcohol-detection accuracies between 91 % and 96 % under controlled conditions [19], [20].

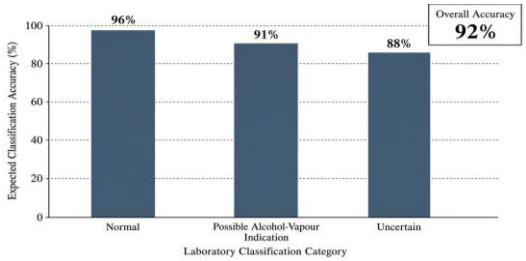


Fig. 5. Expected laboratory classification performance.

Advantages:

- Preliminary in-vehicle alcohol-vapour monitoring
- Real-time warning generation
- Combination of AI, IoT, and embedded technologies
- Modular and expandable design
- Support for research in intelligent transportation systems

Limitations:

- Ambient alcohol vapour detection does not directly measure blood alcohol concentration
- Environmental conditions may affect sensor readings
- False-positive and false-negative readings are possible
- AI performance depends on the quality of training data
- The prototype cannot replace certified breathalyser testing

Multimodal approaches combining alcohol sensing with visual or physiological signals are recommended for future robustness [17], [21].

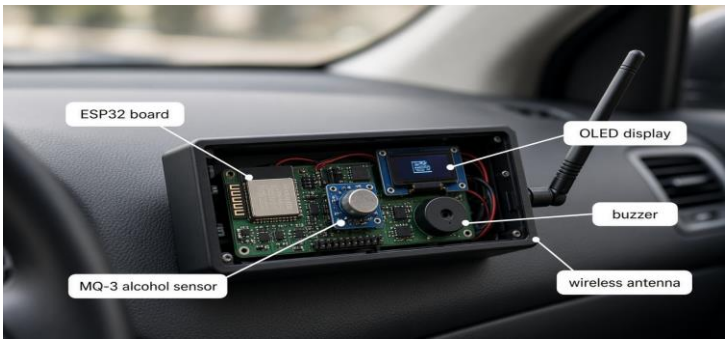


Fig. 6. Proposed prototype model with labelled components (ESP32 board, MQ-3 sensor, OLED display, buzzer, and antenna).

CONCLUSION:

This research proposes an AI-Powered Vehicle-Mounted Intoxication Detection and Driver Safety System to explore intelligent solutions for improving road safety. The system combines alcohol-sensitive sensing, AI-based classification, embedded processing and IoT communication. The proposed architecture provides a foundation for developing a low-cost educational working prototype capable of generating preliminary alcohol-vapour alerts. Although further research, calibration and validation are required, the proposed approach demonstrates how Artificial Intelligence and embedded technologies can contribute to intelligent transportation safety.
Future work will include integration with certified breath-alcohol measurement technology, improved AI models, advanced vehicle monitoring dashboards, integration with fleet-management systems, secure cloud-based event records, and testing under different environmental conditions. A future version may also explore communication with authorised traffic safety monitoring infrastructure, subject to privacy and regulatory requirements.

REFERENCES:

[1] World Health Organisation, Global Status Report on Road Safety 2023, WHO, 2023.

[2] Ministry of Road Transport and Highways, Government of India, Road Accidents in India – 2024, MoRTH.

[3] National Crime Records Bureau, Accidental Deaths & Suicides in India 2024, NCRB.

[4] National Highway Traffic Safety Administration, Alcohol-Impaired Driving statistics for 2024, NHTSA.

[5] M. F. Rahimi and J. Lias, "Alcohol Detection and Engine Cut-Off System using MQ-3 Sensor with GPS and Telegram Alert," Evolution in Electrical and Electronic Engineering, vol. 6, no. 2, 2025.

[6] A. Lightweight In-Vehicle Alcohol Detection Using Smart Sensing and Supervised Learning, Computers, vol. 11, no. 8, p. 121, 2022.

[7] "Real-time Alcohol Detection and Engine Locking System with ESP8266," in Proc. IEEE Conf., 2025.

[8] "Development of Vehicle Alcohol Sensing Engine Lock Safety System," FUOYE Journal of Engineering and Technology, 2026.

[9] "IoT-Enabled Sensor Position Optimisation for Passive In-Vehicle Alcohol Detection and Alert Systems," in Proc. 2025 IEEE International RF and Microwave Conference (RFM), 2025.

[10] "Integrated Driver Safety and Accident Prevention Technology (IDSAPT) using ESP32-CAM and MQ-3," in Proc. IEEE Conf., 2026.

Pavexa – A Porous Concrete Road Solution for Storm Water Management

Satyam Vishwakarma[a], Zarnaab Khan[b], Sanjay Hembram [c]

[a & b] Student of Mechanical Engineering, GGSESTC, Kandra, Chas, Bokaro-827013, Jharkhand

[c] Student of Civil Engineering, GGSESTC, Kandra, Chas, Bokaro-827013, Jharkhand

Abstract: India's roads turn into rivers and lakes every monsoon and cyclone season because of inadequate drainage infrastructure and poor maintenance, disrupting daily life and endangering every section of society, from homeless street dwellers to commuters. Pavexa proposes a porous pavement road system, built from coarse aggregate, fine aggregate, Portland cement and water, engineered so rainwater percolates directly through the road surface into the ground instead of pooling on top. This removes surface water at its source — the root cause of potholes, poor-grip accidents, waterlogged roads and traffic disruption — while simultaneously recharging groundwater. Requiring only correct one-time construction and minimal upkeep thereafter, Pavexa offers municipal corporations a low-maintenance alternative to conventional roads that benefits every economic class equally.

1. Introduction

Every year, India's cities face severe disruption during the monsoon and cyclone season, when heavy and moderate rainfall overwhelms road infrastructure. In the absence of adequate drainage and regular maintenance, roads turn into temporary rivers and lakes, cutting off neighborhoods, damaging vehicles, and putting pedestrians at risk. As urban rainfall intensity continues to rise, this is no longer an occasional inconvenience but a recurring seasonal crisis affecting nearly every Indian city. Pavexa is proposed under the Environment focus area of Idea Tribe, addressing storm water management through a change in road construction itself rather than relying solely on conventional drainage networks.

2. Problem Statement

The core problem is the mismanagement of storm water on Indian roads. Conventional impervious road surfaces do not allow rainwater to pass through them, so all rainfall is forced onto already inadequate and poorly maintained drainage systems. When these systems fail to cope, water accumulates on the road surface itself. This affects everyone — from homeless individuals living on the streets to the wealthiest commuters who rely on these same roads — and existing solutions largely focus on expanding drainage capacity rather than reducing the amount of surface water that needs to be drained in the first place. A more fundamental, low-maintenance solution is required that tackles surface water at its source.

3. Objectives

- To propose a porous pavement road system that significantly reduces surface runoff during heavy rainfall.
- To enable natural groundwater recharge through everyday road infrastructure.
- To reduce monsoon-related hazards such as potholes, waterlogging, poor-grip accidents, and traffic disruption.
- To offer municipal corporations a low-maintenance alternative to conventional road construction.
- To design a solution that is practical and affordable enough for wide urban and semi-urban implementation.

4. Proposed Solution / Innovation

Pavexa is a porous pavement road made from just four ingredients — coarse aggregate, fine aggregate, Portland cement, and water — proportioned to leave interconnected voids within the concrete rather than a solid, impermeable surface. Unlike conventional roads, where numerous problems (potholes, waterlogging, poor grip, traffic disruption) all stem from one root cause — standing surface water — Pavexa removes that root cause entirely by letting the road itself absorb and direct water into the ground. It requires correct construction only once and minimal maintenance thereafter, making it a practical, low-cost long-term alternative to conventional pavement for flood-prone Indian cities.

5. Methodology / Working Principle

When rain falls on Pavexa's surface, its porous structure allows water to be absorbed rather than pooling on top. The water travels down through a layered system — the porous concrete surface, an open-graded base course, and a crushed-stone sub-base that acts as a temporary storage reservoir — before percolating into the natural subgrade, recharging the groundwater table and keeping the surface dry, as shown in Figure 1.

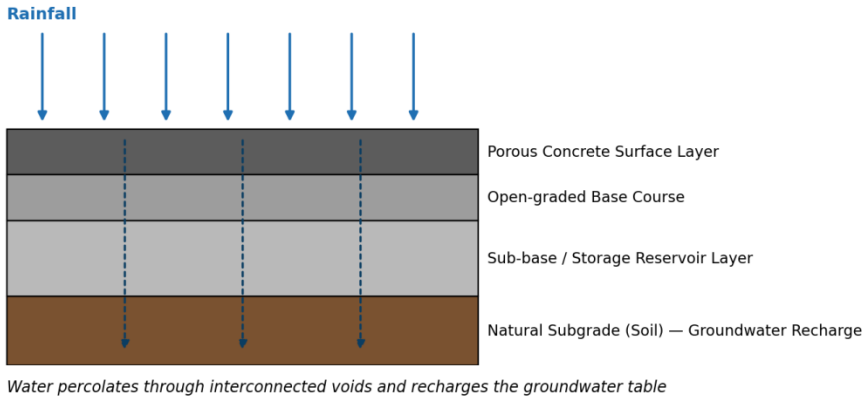


Figure 1 Cross-section of Pavexa showing water infiltration through pavement layers

6. Target Beneficiaries and Expected Impact

Pavexa benefits every section of society that uses Indian roads — from homeless individuals living on the streets to wealthy commuters — since waterlogging affects everyone equally. Municipal corporations benefit directly from lower long-term maintenance costs, as a one-time correct construction replaces recurring drain-clearing and pothole repair. Environmentally, the project supports groundwater recharge in increasingly water-stressed cities. Socially and economically, it reduces monsoon-related accidents caused by poor road grip and standing water, and minimises traffic disruption during the rainy season.

7. Cost Analysis / Estimated Budget / Cost Sheet for Our Ask

The table below presents an item-wise estimated cost for constructing and testing a pilot demonstration stretch of Pavexa (1 Porus block Budget), covering materials, construction, and testing.

S. NO	Item	Specification	Qty	Est. Cost
1.	Concrete (10 – 20 mm)	Field Assesment	500gm	3
2.	Cement	Field Assesment	100gm	1
3.	Dr. Flixit (LH+)	Field Assesment	100ml	2
			Approx Total --	6/-

8. Team & Responsibilities

Team Member	**Role & Responsibilities**
Satyam Vishwakarma	Idea conceptualization, technical development of the porous concrete mix design, and testing (permeability and infiltration trials, quality checks).
Zarnaab Khan	Documentation (report compilation, data recording) design (diagrams, layout, and presentation material).
Sanjay Hembram	design (diagrams, layout, and presentation material).

9. Feasibility and Implementation Plan

Implementation is planned in stages. Phase 1 involves site and soil assessment, including permeability testing and a water-table depth check to confirm suitability. Phase 2 covers mix design and lab trials to balance strength and porosity in the aggregate-cement-water ratio. Phase 3 is construction of a pilot demonstration stretch, followed by Phase 4, field testing under real or simulated rainfall. Phase 5 monitors performance — infiltration rate and load-bearing capacity — across a full monsoon cycle, after which Phase 6 refines the design and prepares a scale-up proposal for municipal bodies. The plan requires basic civil and mechanical engineering expertise, standard construction materials and tools, a partnering municipal corporation for pilot site access, and simple field-testing equipment — all realistically available.

10. Sustainability / Business Model

Pavexa is designed to be funded by city or state government bodies, such as municipal corporations and smart city or urban development schemes, as the primary implementing customer. Its financial sustainability comes from its low-maintenance nature: once correctly constructed, it needs minimal upkeep, saving municipal bodies the recurring costs of drain clearing, pothole repair, and flood damage that conventional roads incur every monsoon. Over time, this cost-saving case can support scale-up funding, and the methodology could also be offered through consultancy or licensing arrangements with construction and infrastructure firms.

11. Future Scope

Beyond the pilot stretch, Pavexa can be scaled to full road networks in flood-prone Indian cities and extended to parking lots, footpaths, residential complexes, and other paved surfaces. Future development could integrate real-time infiltration monitoring through IoT sensors as part of smart city systems, explore ways to increase load-bearing capacity for heavier traffic and highways, and combine the pavement with broader rainwater harvesting infrastructure. The approach can also be adapted and tested across India's varied soil and climate zones for wider deployment.

12. Conclusion

Storm water mismanagement turns Indian roads into rivers and lakes every monsoon, affecting every section of society and straining municipal resources. Pavexa addresses this at its root by using a porous, four-ingredient pavement that absorbs rainwater instead of letting it pool, reducing potholes, accidents, and traffic disruption while recharging groundwater. Requiring only correct one-time construction and minimal subsequent maintenance, it offers a practical, low-cost, and scalable solution that benefits commuters, pedestrians, and municipal corporations alike, making it a strong candidate for pilot implementation and future city-wide adoption.

Dynamic Speed Control of a 3-Phase Squirrel-Cage Induction Motor via Field-Oriented Strategy

Neha Kumari Paul
Department of Electrical and Electronics Engineering
GGSESTC, Bokaro, India
paulneha422@gmail.com

Tanushree
Department of Electrical and Electronics Engineering
GGSESTC, Bokaro, India
kumaritanushree456@gmail.com

Jayanti
Department of Electrical and Electronics Engineering
GGSESTC, Bokaro, India
jayanti.767@gmail.com

Abstract: As in modern world Induction motors are widely used in industrial application and this type of motor are rugged and robust. To change the speed of induction, motor any one parameter either frequency or number of poles are to be changed, now the problem is number of poles are fixed during the design of machine so, cannot change without making any special arrangement. Now the only parameter that remain I frequency but it effect Torque of the machine that depend on the air gap flux, which kept constant to achieve good torque performance at all speed, so to control the speed and torque we FOC or VECTOR CONTROL.
FOC converts the complex stator currents into two orthogonal components, one of which is responsible for speed control, and the other for the electromagnetic torque control, similar to DC machines. FOC has greater stability and capable of fast torque response and accurate speed control. This work explains the concept of FOC and its simulation using Matlab/ Simulink [4].

I. INTRODUCTION:

Field Orientated Control: The Field Orientated Control (FOC) consists of controlling the stator currents represented by a vector. This control is based on projections which transform a three-phase time and speed dependent system into a two co-ordinate (d and q co-ordinates) time invariant system. These projections lead to a structure similar to that of a DC machine control. Field orientated controlled machines need two constants as input references: the torque component (aligned with the q co-ordinate) and the flux component (aligned with d co-ordinate) [1]. As Field Orientated Control is simply based on projections the control structure handles instantaneous electrical quantities. This makes the control accurate in every working operation (steady state and transient) and independent of the limited bandwidth mathematical model. The FOC thus solves the classic scheme problems, in the following ways: • the ease of reaching constant reference (torque component and flux component of the stator current) • the ease of applying direct torque control because in the (d,q) reference frame the expression of the torque is:

$$T \propto \psi_{RS} i_{SQ}$$

By maintaining the amplitude of the rotor flux (ψ_{RS}) at a fixed value we have a linear relationship between torque and torque component (i_{SQ}). We can then control the torque by controlling the

FOC is also known as Vector control technique. It was developed for operating high performance motor smoothly over full speed range[3]. These motors have high dynamic performance which includes fast acceleration and deceleration and capable of fast torque response and accurate speed control. This work explains the concept of FOC and its simulation using Matlab Simulink

In vector model, an AC induction motor is controlled under all conditions. It is like a separately excited DC motor. The field flux linkage and armature flux linkage are created by respective fields. This control generates a three phase PWM motor voltage output.

The output is derived from complex voltage vector to control a complex current vector. The complex current vector is developed by projections or rotations back and forth from motor three phase stator current. The ac drives are considerably more complex than dc drives in case of controlling and estimation.

II. INDUCTION MOTOR MODELING

In Induction Motor Modeling the three-phase supply is converted to two-phase supply. This conversion is done with the help of 'Parks Transformation Matrix'. This machine rated has three HP and Four pole[3].

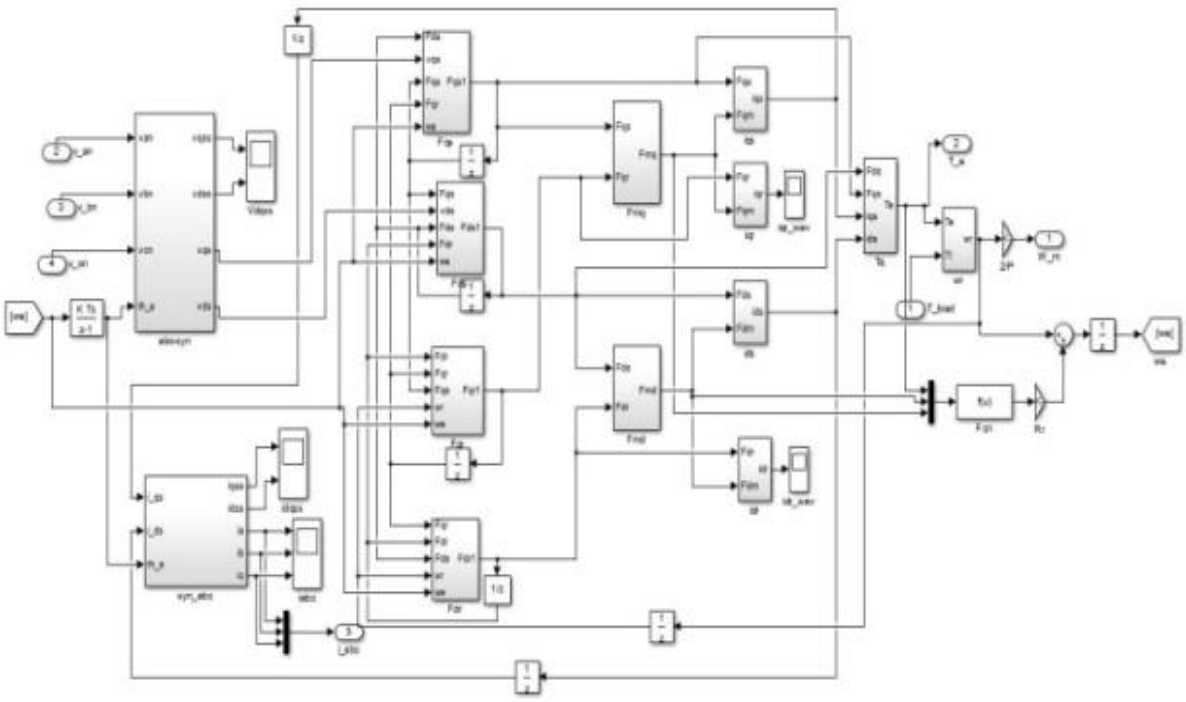


Fig.1. Implementation of induction motor in MATLAB.

III. MODELING OF FIELD-ORIENTED CONTROL

FOC uses current control to control the torque of three phase motors and stepper motors. These motors have high accuracy and bandwidth. FOC can be implemented in either software or hardware.

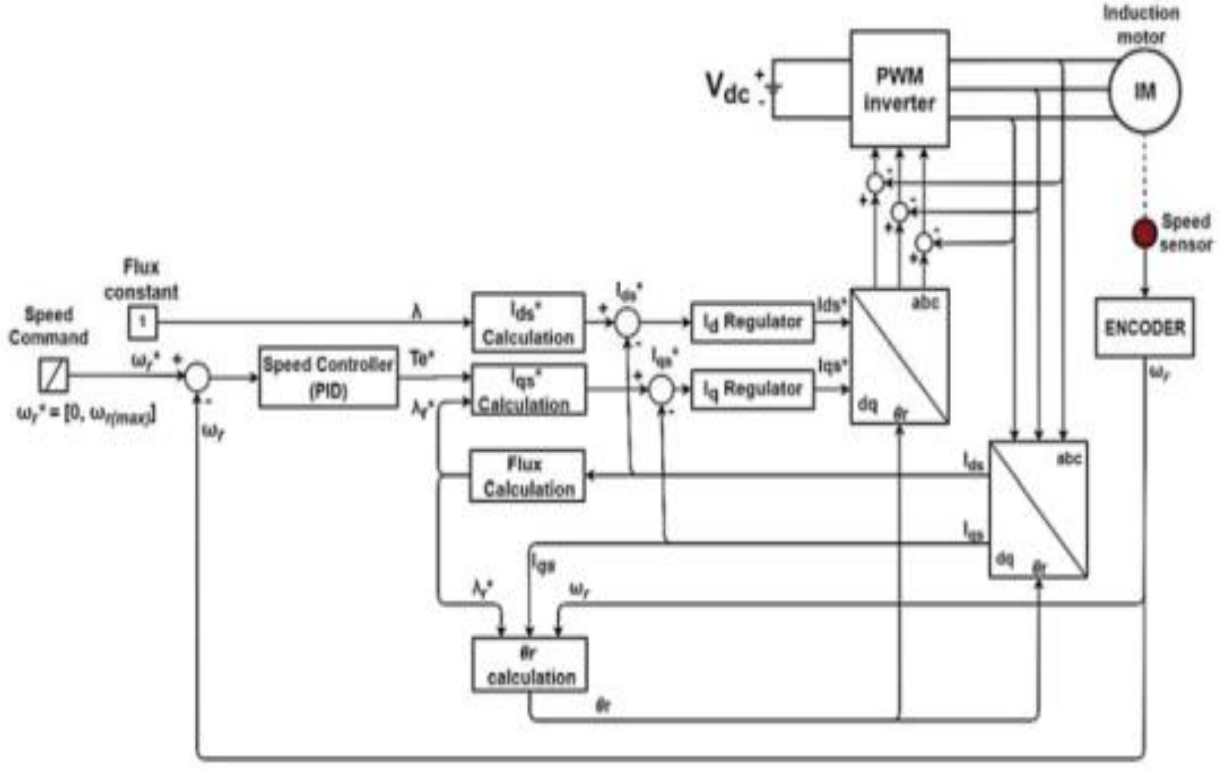


Fig.1. Block diagram of PID speed controller.

The field oriented control structure is based on Clarke and Park transformation matrix.

IV. COMPLETE FIELD ORIENTED CONTROL

This model consists of full model of field-oriented control system. This consists of all connections of inverter, SPWM and FOC system.

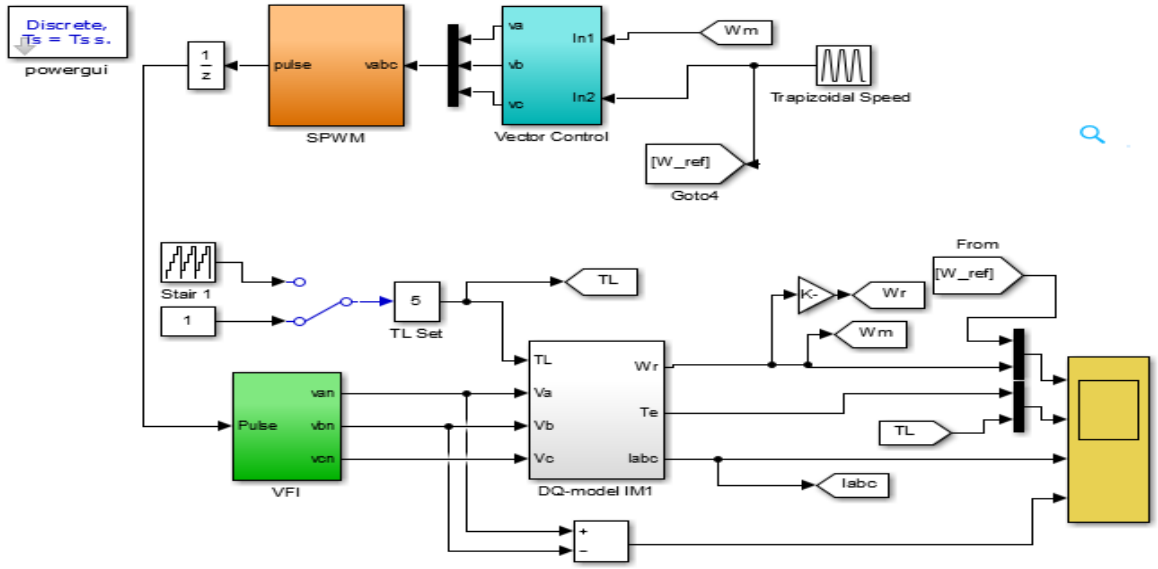
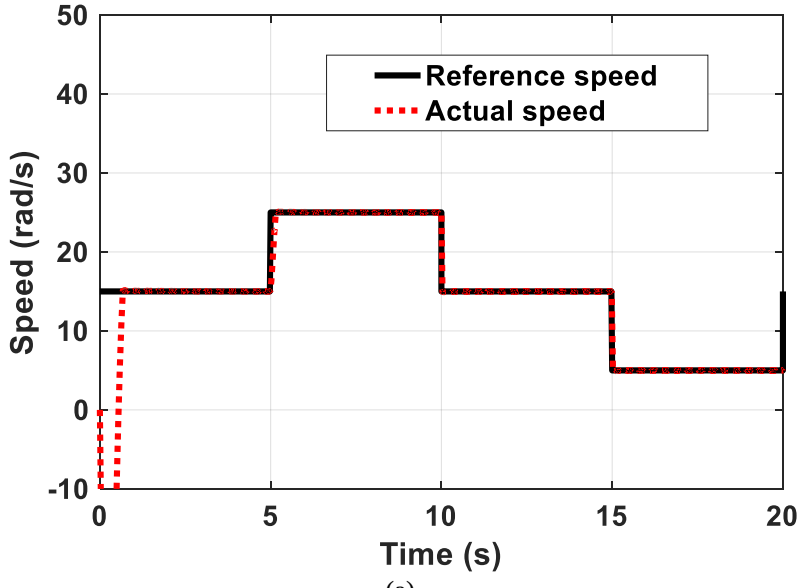


Fig.3 Implementation of Complete FOC Model in MATLAB/SIMULINK.

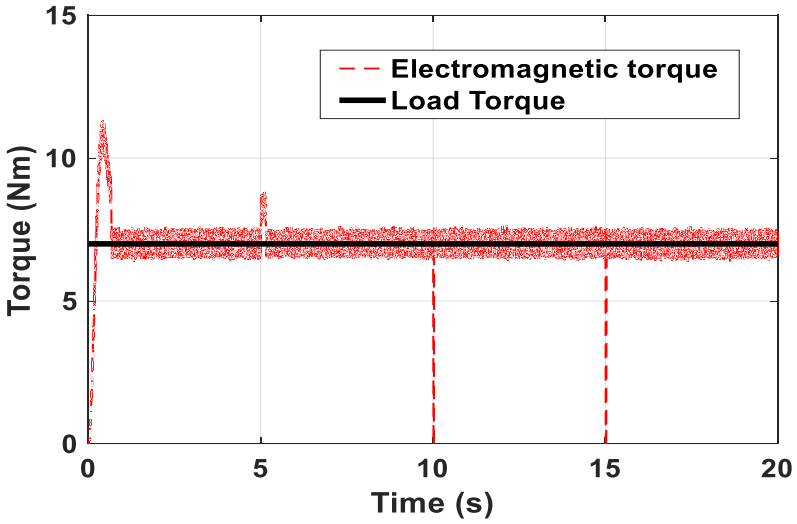
V. SIMULATION RESULTS

The Simulation of Vector Control of Induction Motor is done by using MATLAB/SIMULINK. The below results are shown for speed, torque, stator current at torque load=7,tlim=4, clim=100.

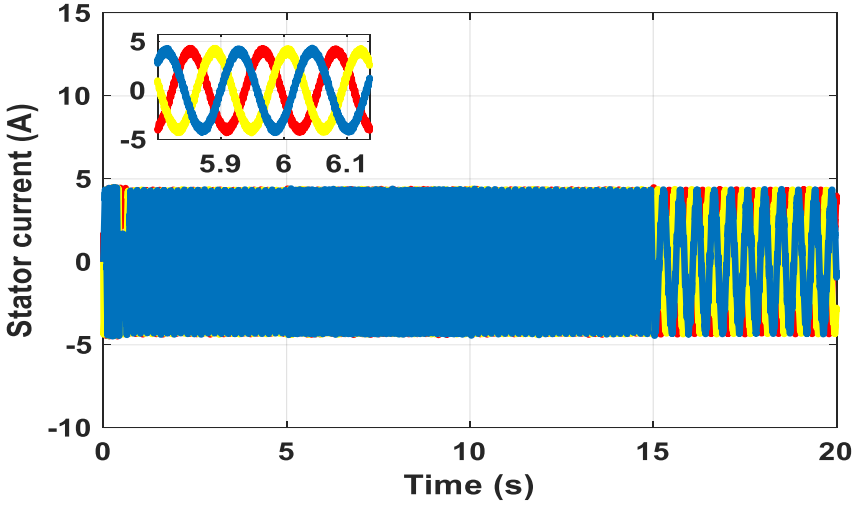
1. Step change in rotor speed at constant torque (T_L= 7 Nm)



(a)



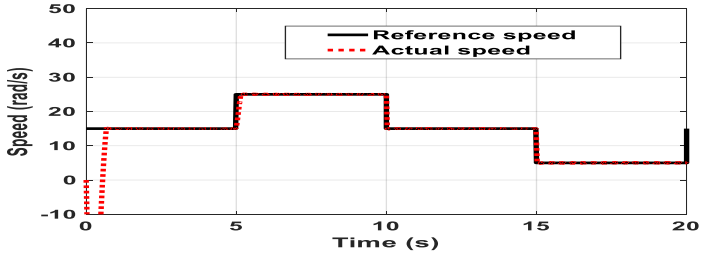
(b)



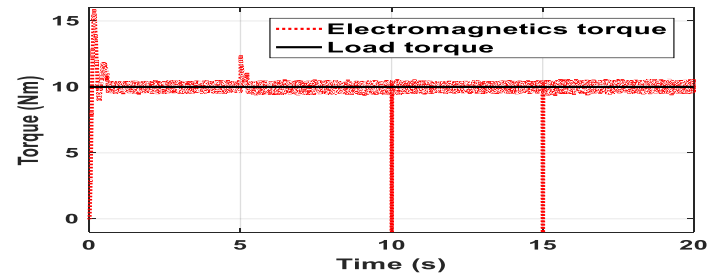
(c)

Fig. 4. Simulation results for step change in rotor speed at load torque of 7 Nm: (a) reference and actual speed, (b) load and electromagnetic torque, and (c) three-phase stator currents.

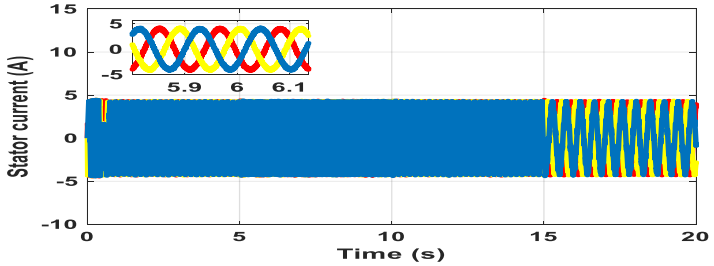
Vector control of induction motor with a 2200W is simulated in MATLAB-SIMULINK. The induction motor used here is four pole, 50 Hz motor. At Torque load =10, Vdc = 560, at T_{lim}= 11.7, C_{lim}=100 for step
2. Step change in rotor speed at constant torque (T_L= 10 Nm)



(a)



(b)

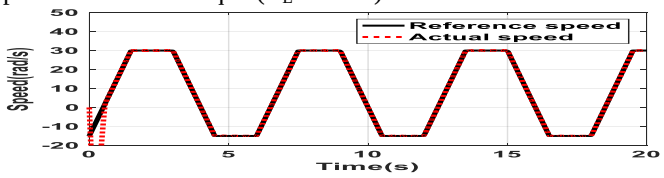


(c)

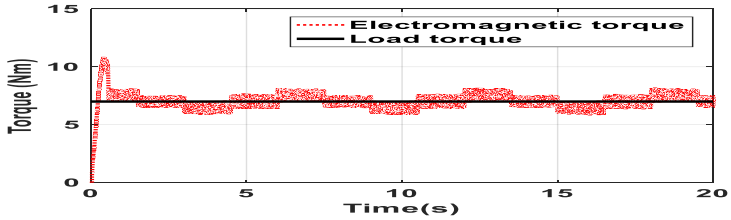
Fig.5. Simulation results for step change in rotor speed at load torque of 10 Nm: (a) reference and actual speed, (b) load and electromagnetic torque, and (c) three-phase stator currents.

Vector control of induction motor with a 2200W is simulated in MATLAB-SIMULINK. The induction motor used here is four pole, 50 Hz motor. At Torque load =7, Vdc = 560, at T_{lim}=5.7, C_{lim}=100 for Ramp

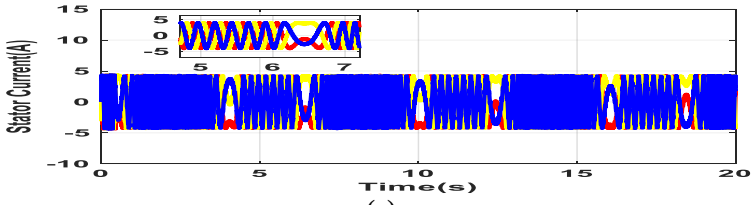
3. Ramp change in rotor speed at constant torque (T_L= 7 Nm)



(a)



(b)



(c)

Fig.5. Simulation results for ramp change in rotor speed at load torque of 7 Nm: (a) reference and actual speed, (b) load and electromagnetic torque, and (c) three-phase stator currents.

VI. CONCLUSION

The Induction Motor drive system, which is based on Field Oriented Control, is developed in this work. By using FOC scheme with induction motor, high starting torque can be obtained like in DC motor. The benefits of FOC are that all the advantages of DC motor can be obtained in FOC-induction motor drive, with the default advantages of

IM over DC motor, which makes the induction motor the most suitable drive for the applications where variable speed operations are necessary with fast response.

VII. REFERENCES

[1] Ion Boldea and Syed A .Nasar, "The Induction Machine Handbook", CRC Press 2002 Second Edition, ISBN: 978-0-8493-0004-2
[2] MahfoozShalaby, "ELECTRIC DRIVES", Dar ElfikrElarabi 2012 Third Edition, ISBN: 977-10-2659-3
[3] MirceaPopescu, "INDUCTION MOTOR MODELLING FOR VECTOR CONTROL PURPOSES", Helsinki University of Technology Espoo 2000.
[4]A. K. Akkarapaka and D. Singh, "The IFOC based speed control of induction motor fed by a high-performance Z-source inverter," in *2014.*
[5]International Conference on Renewable Energy Research and Application (ICRERA), Milwaukee, WI, 2014.

The role of generative ai in enhancing employee productivity and organizational innovation

Sahiba Akram[1], Dr. Shafaque Rahmat[2], Dr. Priyadarshi Jaruhar [3*]

[1]Student, Department of BBA, Guru Gobind Singh Educational Society's Technical Campus, Kandra, Bokaro, Jharkhand
[2]Assistant Professor, Department of MBA, Guru Gobind Singh Educational Society's Technical Campus, Kandra, Bokaro, Jharkhand
[3]***Director,*** Guru Gobind Singh Educational Society's Technical Campus, Kandra, Bokaro, Jharkhand
*Email of the corresponding author: (***Dr. Priyadarshi Jaruhar***) pjaruhar@yahoo.co.in

Abstract

Generative artificial intelligence (GenAI) is changing knowledge work by enabling employees to produce and revise content, analyse information, generate alternatives and develop preliminary solutions with computational assistance. Its organisational significance extends beyond faster task completion: it may alter how knowledge is accessed, how judgement is exercised and how new ideas are developed. This conceptual paper examines the relationship between GenAI, employee productivity and organisational innovation through a focused review of theoretical and empirical literature. It proposes an integrative framework in which GenAI creates value through two pathways: operational augmentation, which supports task execution and output quality, and exploratory augmentation, which expands idea generation, knowledge recombination and experimentation. AI literacy, domain expertise, task–technology fit, employee discretion, managerial support and responsible governance are identified as conditions shaping these pathways. The framework also recognises risks arising from inaccurate outputs, overreliance, confidentiality breaches, bias, intellectual-property uncertainty and convergence in generated ideas. Drawing on the Resource-Based View and the Job Demands–Resources perspective, the paper develops seven propositions for future research. It argues that sustainable gains depend on integrating GenAI into work systems while retaining human judgement for evaluation, contextualisation and accountability.

Keywords: Generative artificial intelligence; employee productivity; organisational innovation; human–AI collaboration; employee creativity; AI literacy; responsible AI.

1. Introduction

Generative artificial intelligence refers to systems that create or transform text, images, code, summaries and other content in response to instructions and contextual information. Its workplace applications now extend across communication, customer service, software development, research, marketing, learning and managerial analysis. Unlike conventional automation focused on predefined routines, GenAI can assist with language-intensive and knowledge-based activities. This makes it relevant not only to operational efficiency but also to the processes through which organisations interpret information and develop new products, services and practices.

Evidence from controlled experiments suggests that GenAI can improve performance on selected professional tasks. Noy and Zhang (2023) found that access to ChatGPT reduced completion time and improved the quality of professional writing tasks in their experimental setting. Brynjolfsson, Li and Raymond (2025) reported productivity gains from a generative AI assistant in customer support, with particularly strong benefits for less experienced workers. These findings do not prove that GenAI improves every job or every dimension of performance. Outcomes depend on task characteristics, employee expertise and the quality of human review.

Innovation presents a more complex question. GenAI can provide alternatives, stimulate brainstorming and make it easier to test early concepts. However, generated suggestions may be inaccurate, derivative or similar across users. The managerial challenge is to create conditions in which AI expands human thinking without replacing critical evaluation or narrowing the diversity of ideas.

2. Research Problem and Objectives

Organisations are adopting GenAI faster than many are redesigning the work practices and governance arrangements needed to use it effectively. Faster drafting or information retrieval may be offset by fact-checking, correction and rework. Similarly, a larger number of AI-generated ideas does not necessarily produce innovation unless employees select, develop and implement useful concepts. Research therefore needs to move beyond simple measures of adoption and examine the mechanisms linking GenAI use with work outcomes.

This paper aims to:

(i) explain how GenAI may influence employee productivity;

(ii) examine its contribution to employee creativity and organizational innovation;

(iii) identify individual and organisational conditions shaping effective human–AI collaboration;

(iv) develop an integrative framework and propositions for future empirical testing.

3. Conceptual Review and Theoretical Foundations

Employee productivity. GenAI may support productivity by reducing time spent on first drafts, routine correspondence, document summaries, code suggestions, information synthesis and preliminary analysis. It can also make specialist knowledge more accessible to employees who have less experience with a task. Noy and Zhang (2023) provide experimental evidence of improvements in time and output quality for professional writing tasks. Brynjolfsson et al. (2025) report that an AI assistant improved the rate at which customer-support agents resolved issues, with larger gains among less experienced employees. These results point to a potential learning and augmentation effect, but productivity should be evaluated through accuracy, usefulness and downstream outcomes as well as speed.

Organisational innovation. Innovation requires more than idea production; it involves identifying opportunities, developing concepts, mobilising resources and implementing solutions. GenAI can

contribute by generating alternatives, connecting information from different knowledge areas, supporting prototypes and accelerating early feedback. Pan (2026) examines GenAI use and employee creativity through creative self-efficacy and creative process engagement, highlighting the relevance of employee-level mechanisms. At the same time, Doshi and Hauser (2024) show that generative AI assistance can improve individual creative output while reducing the diversity of ideas produced across a group. Organisations should therefore consider not only the number of ideas generated but also their originality, variety and implementation value.

The Resource-Based View (RBV) provides a strategic explanation. Barney (1991) argues that advantage can arise from valuable resources that are difficult to imitate. Since widely available GenAI tools may not remain distinctive, organisational advantage is more likely to depend on complementary assets: employee expertise, proprietary knowledge, reliable data, distinctive workflows and the capacity to learn from AI-supported work. The Job Demands–Resources perspective (Bakker & Demerouti, 2007) adds that GenAI can function as a resource by reducing routine effort and improving information access, but may also create demands through verification requirements, learning pressure, uncertainty and intensified performance expectations. A human–AI augmentation perspective treats AI as a complement to human judgement, not as a substitute for responsibility.

4. Proposed Integrative Framework

The framework distinguishes two pathways. The productivity pathway involves GenAI-enabled task assistance, information synthesis, workflow support and feedback. These mechanisms may improve efficiency and output quality when the tool is suitable for the task and employees can evaluate its responses. The innovation pathway involves idea exploration, knowledge recombination, rapid experimentation and cross-functional integration. These mechanisms may support creativity and the development of new products, services and processes when employees retain ownership of the problem and participate in selecting and developing ideas.

The pathways are shaped by six conditions: (1) AI literacy, including the ability to formulate instructions and recognise limitations; (2) domain expertise; (3) task–technology fit; (4) employee autonomy and discretion; (5) managerial support, training and knowledge-sharing routines; and (6) responsible governance covering privacy, security, bias, intellectual property and accountability. Human verification is particularly important where errors may affect customers, employees, finances or safety.

The proposed sequence is: GenAI use → operational and exploratory augmentation → employee productivity and creative process engagement → organisational innovation. AI literacy, task–technology fit, domain knowledge, leadership and governance condition the strength and quality of these relationships.

5. Research Propositions

P1: GenAI use is positively associated with employee productivity when the technology is well matched to task requirements.

P2: Employee AI literacy and domain expertise strengthen the positive relationship between GenAI use and work quality.

P3: GenAI use is positively associated with employee creative process engagement when employees retain discretion over problem framing and idea selection.

P4: Knowledge recombination and experimentation mediate the relationship between exploratory GenAI use and organisational innovation.

P5: Human verification and critical evaluation strengthen the relationship between GenAI-assisted work and reliable task outcomes.

P6: The effect of GenAI on productivity and innovation varies according to task complexity, employee experience and task–technology fit.

P7: Responsible governance and supportive leadership strengthen beneficial outcomes while reducing risks associated with inaccuracy, bias, confidentiality and overreliance.

6. Discussion and Managerial Implications

GenAI should be implemented as part of work-system redesign rather than introduced as an isolated application. Managers should identify tasks where AI assistance is appropriate and distinguish routine, low-risk activities from decisions requiring professional judgement. Employees need practical development in prompt design, source verification, data protection and recognition of unreliable or biased outputs. AI literacy should include critical judgement, not merely tool operation.

Organisations should also decide how productivity gains will be used. If saved time is converted only into higher output expectations, employees may experience work intensification and have little opportunity to undertake higher-value tasks. A developmental approach reallocates capacity towards customer understanding, complex problem-solving, learning and experimentation. For innovation, teams should preserve independent idea generation alongside AI-supported brainstorming so that common model suggestions do not dominate the range of possible solutions.

Performance evaluation should consider quality, accuracy, customer outcomes, learning and implemented innovation rather than counting prompts or generated documents. Governance arrangements should specify approved tools, restrictions on confidential data, review responsibilities, record-keeping and escalation procedures. The OECD (2025) emphasises that GenAI's benefits depend on the task and user experience, and that human expertise remains important. These supports treating GenAI as a capability whose value depends on organisational judgement and complementary skills.

7. Limitations and Future Research

The present study offers a conceptual understanding of the relationship between generative AI, employee productivity and organizational innovation; however, the proposed relationships remain to be validated through empirical investigation. Since the available research largely examines specific professional tasks and selected industry contexts, the findings may vary across occupations, organizational structures and levels of employee expertise. Future studies may therefore explore how

generative AI influences employees performing routine versus knowledge-intensive tasks and whether its impact differs between novice and experienced professionals. Longitudinal research would also be valuable in determining whether productivity improvements are sustained over time, how employees' skills evolve with continued AI use, and whether AI-generated ideas translate into meaningful organizational innovations. Further investigation is needed into the broader organizational implications of generative AI, particularly its influence on employee autonomy, data privacy, algorithmic bias, intellectual-property rights and the diversity of ideas generated through collective human creativity.

8. Conclusion

GenAI can enhance employee productivity by assisting with information-intensive tasks and can support organisational innovation by widening exploration, enabling knowledge recombination and accelerating experimentation. These benefits are not automatic: they depend on task suitability, employee expertise, AI literacy, human verification, managerial support and responsible governance. This paper contributes an integrative framework that distinguishes operational augmentation from exploratory augmentation and explains how both may translate into organisational value. Organisations are most likely to benefit when GenAI strengthens rather than displaces human judgement, preserves employee ownership of work and is embedded in accountable, learning-oriented work systems.

References

Bakker, A. B., & Demerouti, E. (2007). The Job Demands–Resources model: State of the art. Journal of Managerial Psychology, 22(3), 309–328. https://doi.org/10.1108/02683940710733115

Barney, J. (1991). Firm resources and sustained competitive advantage. Journal of Management, 17(1), 99–120. https://doi.org/10.1177/014920639101700108

Brynjolfsson, E., Li, D., & Raymond, L. R. (2025). Generative AI at work. The Quarterly Journal of Economics, 140(2), 889–942. https://doi.org/10.1093/qje/qjae044

Calvino, F., Reijerink, J., & Samek, L. (2025). The effects of generative AI on productivity, innovation and entrepreneurship. OECD Artificial Intelligence Papers, No. 39. OECD Publishing. https://doi.org/10.1787/b21df222-en

Doshi, A. R., & Hauser, O. P. (2024). Generative AI enhances individual creativity but reduces the collective diversity of novel content. Science Advances, 10(28), eadn5290. https://doi.org/10.1126/sciadv.adn5290

Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. Science, 381(6654), 187–192. https://doi.org/10.1126/science.adh2586

Pan, Y. (2026). Exploring the relationship between generative AI usage and employee creativity: A dual-pathway mediation model. European Journal of Innovation Management, 29(3), 799–823. https://doi.org/10.1108/EJIM-02-2024-0213

A Token of Appreciation

The Organizing Committee of the **3rd ISTE Student Convention** at Guru Gobind Singh Educational Society's Technical Campus expresses its heartfelt gratitude and sincere appreciation to **Mr. Ankur Mandal, Proprietor, Anirban Udyog**, for his generous sponsorship and valuable support towards the successful organization of the convention held on 18–19 November 2026.

Your generosity, encouragement, and commitment to promoting technical education reflect your deep sense of social responsibility and dedication to nurturing young talent. Your valuable contribution has helped create an inspiring platform for students, researchers, and academicians to exchange knowledge, showcase innovations, and pursue excellence.

Your support is more than a contribution; it is a gesture of encouragement that inspires young minds to dream bigger, innovate further, and achieve greater heights.

We sincerely acknowledge your association with this prestigious event and wish you and Anirban Udyog continued success, prosperity, and excellence.

With sincere gratitude and highest appreciation.

Organizing Committee

3rd ISTE Student Convention

ISTE Student Chapter

Guru Gobind Singh Educational Society's Technical Campus

TEXRICA™
Texrica India Organics Private Limited

A Promise For Yarn
ART - YARN EXPORTS (INDIA) PVT. LTD.

VARTMAN™
TEXTILES (FZE)

Proudly Honoured with

MATEXIL Export Awards 2024–25

GOLD AWARD
Woman Entrepreneur

BRONZE AWARD
Export Excellence



Art-Yarn celebrates two MATEXIL Export Awards 2024–25:
Gold for Woman Entrepreneur and ***Bronze for Export Excellence.***

A proud milestone made possible by our people and partners.

GGSESTC

B++ NAAC

GURU GOBIND SINGH EDUCATIONAL SOCIETY'S TECHNICAL CAMPUS

COLLEGE OF ENGINEERING AND MANAGEMENT

Approved by AICTE, Ministry of Education, Govt. of India, New Delhi, Affiliated to Jharkhand University of Technology, Ranchi

NH-18, KANDRA (PURULIA ROAD) CHAS, BOKARO, JHARKHAND - 827013

A **Fifth Generation Engineering & Management College** with **NAAC Accreditation**



B.Tech.
CSE, CSE (DS), CSE (CS)
Fashion Technology, ME, CE, EEE, ECE

B.B.A. & B.C.A.
For BBA, Passed 10+2 with Minimum 45% Marks for General & 40% for Reserved category in any discipline Science/Arts/Commerce

M.B.A.
Finance, Human Resources (HR) Marketing

Diploma In Engineering
EEE, ME, Civil (For Diploma 10th Complete from any branch)

Best results among JUT affiliated Engg. & Management Colleges

Highest Package 14 LPA

UPTO 100% SCHOLARSHIP

ISO 9001:2015 EAT RIGHT CAMPUS CERTIFIED

SCHOLARSHIP
(Merit Cum Means based)
SAT-2027

ISTE Best Faculty Convention Award 2025

- No Tuition Fees for E-kalyan eligible Students.
- Manki Munda Scholarship by Govt. of Jharkhand for Girls students of open category.
- Flexible time table for Working Professionals in B.Tech. & M.B.A. as a regular course.
- Excellent Placement Assistance.
- NIRF Ranking-MOUs with Industries
- Gold Medalists from GGSESTC

NO. 01 COLLEGE OF ENGINEERING, MANAGEMENT, APPLICATION & FASHION IN JHARKHAND

BUILD YOUR FUTURE WITH EXCELLENCE



NOTE : PLEASE REFER OUR WEBSITE FOR MORE DETAILS

www.ggsestc.ac.in | info@ggsestc.ac.in

Follow us:

06542-265293 | 7992284995
7992415822 | 9822732264

Prof (Dr.) Priyadarshi Jaruhar,
MTech, PhD (NIT J), SLM ISTE
Director, GGSESTC, Bokaro