Advanced International Journal for Research (AIJFR)
E-ISSN: 3048-7641 ● Website: www.aijfr.com ● Email: editor@aijfr.com

Strengthening Digital Payment and Security using AI, Big Data, and Cybersecurity in India

Jignesh Bari[1], Ishita Holkar[2], Ritesh Pawar[3], Rutuja Dekate[4]

[1,2,3,4]Student, Department of Management, IIEBM Indus Business School, Pune

Abstract

India's digital payment system has grown fast over the last ten years. The Unified Payments Interface and services linked to Aadhaar have helped hundreds of millions of people live in a cashless economy. More people are using their smartphones now, which is why. I believe this is a good deal. This growth is real and very important. We should pay attention to India's digital payment system. The risks have also grown as more people use payments. There have been times when people have lied and stolen data. Attacks that people didn't expect have cost them money. This is a problem. This paper examines the enhancement of India's digital payment system through the integration of artificial intelligence, big data analytics, and robust cybersecurity measures.

We learned about the threats by looking at research, government reports, and data. We also looked at technologies and the environment and found areas where things could be better. Our main findings show that fraud detection systems that use AI are better than those that use rules. Biometric authentication is very useful. Behavioural authentication gives you more security. More and more institutions are using zero-trust architecture as a standard.

The technology isn't enough. Social engineering attacks are effective due to users' lack of awareness regarding the threat. Designing for people who are digitally literate and making it easier to file complaints are just as important as any algorithm. The paper concludes with suggestions for policymakers, banks, tech companies, and civil society. They are all responsible for making sure that a payment system that affects more than a billion people in India is safe. India's digital payment system and artificial intelligence will help keep it safe and secure. Everyone should be able to trust and use India's digital payment system safely. Digital payments and AI in India are here to stay. Will keep getting bigger. We need to get used to it.

Keywords: Digital Payments, UPI, Artificial Intelligence, Big Data, Cybersecurity, Fraud Detection, Financial Inclusion, India, Zero-Trust Architecture, Biometric Authentication.

1. Introduction

Over the past ten years, India has changed the way they deal with money in a big way. There wasn't just one thing that happened. There were many things. For example, when they started Aadhaar in 2009, when they stopped using cash in November 2016, and when they added the Unified Payments Interface that same year. That all together slowly, then quickly, got hundreds of millions of people to use digital

payments. UPI was processing more than 13 billion transactions every month by 2024. These numbers are very surprising. There is a person behind every number. Like a street vendor who uses a QR code, a student who pays their hostel fees with their phone, or a worker who sends money home without having to go to the bank.

This is a story that will make you very happy. It's a big deal to get a lot of people to use digital payments. There is a bad side. People who join a payment platform for the first time are also at risk of being targeted by criminals. These bad people have been following the cash. Trying to fool people. The Indian Cybercrime Coordination Center got more than 1.4 million reports of cybercrime in 2023. Many of them were about fraud with digital payments. A lot of people lost money. And the tricks these bad people used weren't hard to figure out. They just called people, sent fake links, or used fake QR codes. They just had to find someone who didn't know what to watch out for.

It is this question that will be attempted to answer through this paper. This is not about whether digital transactions are positive in themselves. This is obvious. The question is on whether the framework within which digital payments can be safeguarded is sufficient to prevent those who intend mischief from succeeding in their plans. AI systems are efficient in detecting fraudulent activity and identifying malicious behaviour in a very short period of time. In a matter of a single second, they are able to analyse millions of transactions. Big data analysis helps to sort out the mess of information. Cybersecurity frameworks like encryption, tokenisation, and zero-trust architecture play a critical role in securing data and finances.

The data used in this paper will come from various sources. Sources such as academic papers, official reports, and statistics, for instance. This will ensure that a comprehensive analysis of India's progress in terms of payment security is achieved. The areas covered will include successes, failures, and future improvements. This paper will be organised using the following format. Background information, methodology, analysis, results, and recommendations, respectively. The objective of this paper is not to instill fear in anyone but to inform everyone of what is happening in India concerning payment security.

1.1 Research Problem:

In most of the research published on digital payments security in India, technology and policy have remained independent spheres of research. In the literature that focuses on artificial intelligence fraud prevention, little has been written on the regulatory framework under which the AI is employed. In the literature on user behaviour, no attention has been paid to the intricacies of the attack mechanisms being discussed. In the literature on rural vulnerability, no emphasis has been placed on architectural modifications needed to mitigate the problem. There has been progress in many areas, but not in the integration of those advances.

Another major flaw in much of the research in this area is the presumption that the problem of security in digital payments is a technical one. The body of evidence accumulated through this literature review proves the exact opposite – most fraud cases exploit human vulnerabilities rather than system vulnerabilities. A research approach focused on the technical aspect of security without addressing the user experience will not yield optimal results. The aim of this paper is to address both problems simultaneously.

1.2 Problem Statement:

The Indian digital payment ecosystem is vast and thriving. However, the security measures protecting it are inconsistent across different parts of the country. The legislation governing it struggles to keep up with emerging issues. Not all of the users can use digital payments securely. Some are even facing losses due to fraud. The population's confidence in it is very low. Moreover, those who require digital payments the most, such as rural inhabitants, new smartphone users, and the uneducated population, are the most vulnerable targets. Digital payments are utilised in India by the population. Therefore, this paper aims to determine the necessary actions required to ensure the safety of the Indian digital payment ecosystem for all payment users. In other words, it seeks to identify which combination of technological improvements, regulations, and public education on digital payments is needed for ensuring the safety of digital payments for all users.

2. Literature Review

The research on payment security in India is one area that brings together computer science, behavioural economics, law, and public policy. Numerous works have been written on the subject since UPI launched. In the light of all that has been discussed above, certain concepts emerge from these publications. Firstly, digital payment security in India is a bargain. The second thing about digital payment security in India is that it involves intelligence in the identification of vulnerabilities, although this cannot be the only solution. As it is seen from discussions on payment security in India, one thing that always turns out to be the weak link is humans themselves. On top of that, the Government of India has developed certain regulations to facilitate payment security in India, which do not appear to be sufficient for addressing emerging challenges. Finally, there appears to be a rural/urban disparity in digital payment security in India. Below are going to be considered concepts related to payment security in India and identified by scholars.

2.1 The Growth and Vulnerability of India's Digital Payment System

According to Arunachalam (2024), the Indian digital trust revolution led to the integration of citizens in the banking system but resulted in creating a massive vulnerable zone. Specifically, he analysed the statistics between 2022 and 2025. Identified that the number of transactions conducted via UPI increased by around 50 per cent annually. Although numerous small banks and payment services were unable to implement mechanisms that would detect fraud. Thus, there was a security gap that lasted for two to three years which criminals utilised effortlessly.

He identified that approximately 3.6 million incidents of cyberfraud occurred within the scope of the timeframe. Such attacks are predicted to result in losses amounting to 1,200 billion rupees by 2025. In other words, the threat is currently relevant. Shahu (2025) reviews the situation too. But unlike the previous author, he investigates the correlation between AI usage and compliance with UPI regulations.

And his research highlights an essential issue in the realm of Indian payments:

* The openness of UPI makes it successful but hard to secure.
* A system made for access does not work well with the extra checks that good security needs.

Shahu thinks AI can solve this problem. It can provide security without causing delays or complexity that would make users leave. UPI and AI are connected in this solution. The goal is to make UPI secure and easy to use.

2.2 Artificial Intelligence in Fraud Detection

Keerthna et al. (2024) tested AI-powered anomaly detection against rule-based benchmarks on a dataset of UPI transactions and found that ensemble machine learning models—specifically combinations of Gradient Boosting and Random Forest classifiers—achieved fraud detection accuracy above 94 per cent, compared to 71 per cent for rule-based systems. The improvement is not marginal. It represents a meaningful reduction in successful fraud, and it is achieved in real time, without adding perceptible delay to legitimate transactions. The study also highlighted the challenge of class imbalance — fraudulent transactions make up less than 0.1 per cent of all transactions — and showed that SMOTE oversampling and cost-sensitive learning effectively addressed this problem.

Mukkamala (2023) focused specifically on recurrent neural networks, arguing that the sequential nature of transaction data makes temporal models uniquely suited to payment fraud detection. A fraudster who takes over an account does not simply make one large transfer — they typically mimic the account holder's behaviour before making their move. An LSTM network that has learned the rhythm of a user's transactions will detect the deviation from that rhythm even if the individual transaction looks superficially normal. Mukkamala's model achieved an F1-score of 0.96 on a benchmark fraud dataset, a result that compares favourably with anything published in this field.

Padmakar et al. (2024) extended the AI fraud detection framework to include contextual signals – device fingerprinting, geolocation patterns, time-of-day behaviour, merchant category codes, and social network indicators – arguing that no single transaction should be evaluated in isolation. Their concept of "360-degree user risk profiling" treats each transaction as a data point in a much larger behavioural story, and anomalies become visible only when the full story is visible. This approach is computationally intensive but increasingly practical as processing costs fall and cloud-based analytics become standard infrastructure for even mid-sized financial institutions.

Priya et al. (2025) tackled what may be the hardest practical problem in AI-based fraud detection: latency. UPI is designed for sub-second transaction completion. Any fraud detection system that adds more than a fraction of a second to that process will generate user complaints and abandoned transactions. Their solution — pre-computing account-level risk scores asynchronously while reserving real-time inference only for transaction-level features — achieved the detection performance of a full model at a fraction of the processing delay. This architectural innovation may prove as important as any algorithmic advance.

Singh et al. (2025) explored the frontier of AI applied to IoT-connected payment devices — smart speakers, wearables, and connected point-of-sale terminals. Each of these devices generates behavioural data that, when fed into an AI fraud detection system, creates a multi-dimensional authentication context that is very difficult for an attacker to replicate. A fraudster who has stolen a UPI PIN still cannot replicate the specific pressure pattern of the victim's fingerprint on a particular device or the specific rhythm with which that person typically navigates a payment app. The convergence of IoT and AI represents the next generation of payment security, and India is beginning to see its first deployments.

2.3 Big Data Analytics and Biometric Authentication

Patra et al. (2022) created and examined an integrated framework which included artificial intelligence, large-scale data processing, and biometrics for authentication purposes. As a result, it was found that using additional authentication via fingerprints, irises, or faces together with big data decreased the percentage of unauthorised transactions by 78 per cent in comparison with transactions using only personal identification numbers within a pilot study. The Big Data dimension becomes significant because of the complex verification of biometric data (comparison with a template), which should be done quickly despite possible variances in lighting, angles and pressure. Apache Kafka for stream processing and Apache Spark for large-scale analytics provided the infrastructure backbone that made this possible at scale.

Das (2024) documents a different but equally important application of big data in Indian finance: The concept of alternative credit scoring implies the aggregation and analysis of data on how users behave on mobile devices. In particular, their internet surfing habits, e-commerce transaction history, regularity of utility payments, and other behavioural features can help to develop a comprehensive picture of a person's creditworthiness. As estimated by Das, alternative credit scoring allows giving credit to more than 190 million people without traditional bank histories. Security-wise, the idea of having more data means getting better fraud detection instruments.

KAGE et al. (2023) raise an important counterpoint. Their study of cybersecurity in rural digital transactions found that the same big data collection practices that power fraud detection raise serious privacy concerns among users who are not fully aware of what is being collected or how it is used. The Supreme Court's 2017 Puttaswamy judgement established privacy as a fundamental right, and the Digital Personal Data Protection Act 2023 operationalised that right in the payment context. KAGE et al. argue that informed, granular consent – as embodied in the account aggregator framework – is not just a legal requirement but a trust requirement. Users who do not trust a system will avoid it, and a payment system that is avoided is a payment system that has failed.

2.4 Cybersecurity Frameworks and Technologies

Saini (2023) provides a comprehensive review of the technical security infrastructure underpinning electronic payment systems. On encryption, the study confirms that AES-256 has become the standard for data in transit and at rest, with RSA-2048 or ECC-256 used for key exchange and digital signatures. On tokenisation, Saini documents the RBI's October 2022 mandate requiring all card-on-file transactions to use tokenised card data, eliminating the need for merchants to store raw card numbers. This single regulatory intervention removed one of the most common sources of large-scale card data theft at a stroke.

Zero-Trust Architecture — the principle of "never trust, always verify" — has moved from a theoretical framework to an operational standard in India's larger financial institutions. Arunachalam (2024) credits the RBI's 2021 Master Direction on IT Governance with accelerating this shift, requiring continuous verification of every user, device, and network flow regardless of location. The practical components of zero-trust — microsegmentation, identity-centric access control, and continuous behavioural monitoring — significantly raise the cost of both external attacks and insider threats.

Goli et al. (2024) document the evolution of security operations centres in Indian banks, showing how AI augmentation has transformed what was once a largely human-driven monitoring function into a system

capable of correlating signals across millions of events per second. The SIEM platforms now deployed at major banks can identify complex, multi-stage attacks that would be invisible to individual rule-based detectors, dramatically reducing the dwell time of undetected intrusions.

Soundenkar et al. (2024) place these technical developments in a governance context, arguing that AI-powered risk management must be paired with clear accountability structures. A fraud detection system that flags a legitimate transaction must provide a coherent explanation for that decision. A system that misses a fraud must be capable of being audited to understand why. Explainability is not just a user requirement — it is increasingly a regulatory one.

2.5 Blockchain and Distributed Ledger Technology

Several researchers have explored blockchain as an alternative or supplement to centralised security architectures in payment systems. Saini (2023) and Das (2024) both analyse blockchain's properties – cryptographic immutability, decentralised consensus, and transparent auditability – as attractive features for tamper-resistant transaction records and decentralised digital identity management. The Account Aggregator Framework, though not technically a blockchain, embodies the same principle of user-controlled, consent-based data sharing that blockchain proponents advocate.

The practical limitations are real. Permissionless public blockchains cannot handle the transaction volumes that UPI processes — tens of millions of transactions daily — without sacrificing either security or speed. The literature points toward hybrid permissioned architectures, such as Hyperledger Fabric, as the more viable path for regulated financial environments. Das (2024) also notes the emerging combination of AI and blockchain, where AI analyses blockchain transaction patterns for anomalies while blockchain provides an immutable audit trail for AI model decisions – an architecture that addresses both fraud detection and explainability requirements simultaneously.

2.6 Social Engineering, User Behaviour, and Trust

Sharma et al. (2024) make a finding that deserves to sit at the centre of any serious discussion of digital payment security in India: social engineering attacks account for approximately 70 to 75 per cent of digital payment fraud cases. The figure is not a rounding error. Three in four fraud cases succeed, not because a system was hacked but because a person was deceived. The specific mechanics have been well documented. A fraudster, typically impersonating a bank official or NPCI representative, contacts a victim and creates urgency — "Your account will be blocked" and "There is a pending KYC update." The victim is then guided to approve a UPI Collect Request, believing they are receiving a refund or a verification payment. The PIN they enter authorises a debit, not a credit. The attack requires no technical skills. It requires only a victim who does not know that receiving money on UPI never requires a PIN.

Aljaradat et al. (2023) examine trust formation using an extended UTAUT model that incorporates perceived cybersecurity risk and digital trust as mediating variables. Their structural equation modelling found that perceived cybersecurity risk carries the largest negative effect on digital trust among all modelled variables ($\beta = -0.47$, $p < 0.001$) and that trust mediates 64 per cent of the total effect of cybersecurity risk perception on behavioural intention. In plain terms: if users believe digital payments are risky, they will not use them, regardless of how good the underlying technology is. Trust is not a soft, secondary concern — it is the load-bearing variable in the adoption model.

Yamuna (2024) adds a nuance that is both encouraging and cautionary. Her study found that users who were simply told that AI-powered security was operating on their transactions reported average perceived safety scores 23 per cent higher than users who received no such information, even when both groups had identical actual security protection. The "AI assurance effect" is real — transparent communication about security technology builds trust independently of technical literacy. The cautionary note is that this effect cuts both ways: a well-publicised fraud incident can destroy trust rapidly, even if the fraud affected only a tiny fraction of users.

Jyothi (2024) examines the legal dimensions of social engineering fraud, arguing that India's existing cybercrime framework — rooted in the IT Act 2000 and its 2008 amendments — is structurally ill-suited to prosecute AI-powered fraud where no single identifiable human actor is directly responsible for each step in an automated attack sequence. The DPDP Act 2023 strengthened data protection rights but did not close this prosecution gap. Jyothi recommends dedicated AI-specific legislation that assigns civil and criminal liability with clarity.

2.7 Machine Learning Models — A Comparative View

The academic literature presents a range of machine learning architectures applied to payment fraud detection, each with distinct strengths and limitations. Logistic regression remains useful as a fast, interpretable baseline but struggles with the non-linear patterns that characterise modern fraud. Random Forest handles class imbalance well and operates comfortably in batch-scoring environments. Gradient boosting — particularly XGBoost — delivers the highest precision on structured tabular data and is widely deployed for daily fraud report generation. LSTM networks capture the sequential nature of transaction behaviour and excel at detecting account takeover fraud, though they are computationally intensive to train. Autoencoder networks, operating without labelled fraud data, are the most effective tool for detecting entirely novel fraud patterns that supervised models have never seen. Graph neural networks, the newest entrant, are designed specifically to identify fraud rings and money mule networks by analysing the relationship structure between accounts rather than individual transaction features.

No single model dominates across all contexts. Production deployments at leading Indian payment processors typically use ensemble approaches — combining the speed of gradient boosting for routine transactions with the sequential sensitivity of LSTMs for flagged accounts and the novelty detection capability of autoencoders for emerging threats (Priya et al., 2025; Keerthna et al., 2024).

2.8 Regulatory and Legal Frameworks

Jyothi (2024) provides a systematic legal analysis of India's electronic banking framework. The IT Act 2000 established legal recognition for electronic contracts and created the first cybercrime prosecution framework. Section 66 addressed unauthorised computer access, Section 66C covered identity theft, and Section 66D criminalised cheating by impersonation using computer resources. The 2008 amendments added Section 43A (compensation for data protection failures) and Section 72A (penalties for unlawful disclosure of information). These provisions have supported numerous prosecutions, but their architecture assumes a human perpetrator at each stage — an assumption that AI-powered fraud increasingly violates.

Raghib et al. (2024) analyse the Digital Personal Data Protection Act 2023 as a landmark improvement in India's data governance landscape. The Act establishes the right to information about data processing, the right to correction and erasure, mandatory breach notification within 72 hours, and a Data Protection Board

empowered to award compensation up to ₹250 crore for significant data breaches. For payment security, the most significant provision is the extraterritorial clause — the act covers any entity that processes personal data of Indian residents, regardless of where the entity is incorporated. This substantially expands the compliance obligations of global payment processors operating in India.

Arunachalam (2024) and Shahu (2025) both highlight the RBI's regulatory sandbox as an important innovation accelerator. By allowing early-stage companies to test AI-driven fraud detection, blockchain-based KYC, and federated learning models in a supervised low-compliance environment, the sandbox has shortened the path from research to production deployment and given regulators first-hand experience with technologies they may subsequently need to regulate.

3. RESEARCH METHODOLOGY

3.1 Objectives:

1. The research aims to evaluate the effectiveness of AI and machine learning techniques in detecting online digital payment fraud in India.
2. The researcher will explore the role of big data analytics and biometrics in securing online payments in India.
3. The researcher will analyse the security mechanisms employed by the Indian banks and other financial organisations, including zero trust architecture, tokenisation, encryption, and security operation centres.
4. The researcher will investigate how behavioural factors and social engineering contribute to online digital payment frauds and the effect of digital literacy on payment security in India.
5. The researcher will verify whether existing laws and policies in India, including the IT Act 2000, DPDP Act 2023, and guidelines from RBI/NPCI, are adequate to address cybersecurity risks in the future.
6. The researcher will identify gaps in technology, regulations, user knowledge, and design that make online digital payments in India vulnerable.
7. The researcher will provide actionable recommendations supported by evidence to the Indian government, banks, technology companies, and non-governmental organisations regarding how to improve the security of online digital payments in India. I will provide these recommendations for AI and machine learning technologies and digital payment systems.

3.2 Nature and Sources of Data

In terms of data collection, the research has made use of only secondary sources. No primary survey was carried out. It would be appropriate at this point to explain why there is no primary survey in the research. The reason for this is the scope of the problem under investigation, which includes not only technological but also legal, behavioural, and public policy issues.

3.2.1 Secondary Data

Secondary data was collected from:

- Peer-reviewed academic journals published on IEEE Xplore, Google Scholar, SciSpace, SSRN, and ACM Digital Library
- Conference proceedings from IEEE GINOTECH 2025 and related fintech and cybersecurity conferences
- Annual reports and circulars from the Reserve Bank of India (RBI) and National Payments Corporation of India (NPCI)
- Reports from the Indian Cybercrime Coordination Centre (I4C), Ministry of Home Affairs
- Legislative texts, including the Information Technology Act 2000, the IT Amendment Act 2008, and the Digital Personal Data Protection Act 2023
- Published theses and working papers from recognised research institutions
- Institutional data on UPI transaction volumes from NPCI's publicly available statistics portal

3.3 Inclusion and Exclusion Criteria

The literature review had a systematic methodology for selecting sources. Sources were chosen based on the following criteria: publication years 2016-2025; topic of research related to digital payment security, application of artificial intelligence to fraud detection, or other cybersecurity topics pertaining to the Indian environment; peer-reviewed; and conducted by an institutionally authoritative body. Those sources published prior to 2016 that did not discuss basic legal frameworks were not considered relevant. In addition, sources that only discussed payment systems unrelated to India were omitted.

3.4 Analytical Approach

The analysis will adopt a thematic synthesis methodology. The published findings have been structured around the themes that emerged from the literature review: application of AI and ML in fraud detection; big data and biometric methods; cybersecurity framework; social engineering and human behaviour; regulatory environment; and gap analysis. In relation to each theme, published findings have been analysed to detect any points of agreement among different findings, conflicting findings, and areas of limited evidence. The gaps analysis and recommendations sections will directly benefit from this thematic synthesis.

3.5 Hypothesis

H_1 (Technology Effectiveness): AI-powered anomaly detection and machine learning-based fraud classification significantly reduce the rate of successful fraudulent digital payment transactions compared to legacy rule-based systems.

- **H_0 (Null): There is no significant difference in fraud detection outcomes between AI-powered and rule-based systems.**
- **H_1 (Alternative): AI-powered systems achieve significantly higher fraud detection accuracy and lower false negative rates than rule-based systems.**

H_2 (Trust and Risk Perception): **Perceived cybersecurity risk has a significant negative effect on user trust in digital payment systems, which in turn significantly reduces behavioural intention to use those systems.**

- **H_0 (Null): Perceived cybersecurity risk has no significant relationship with user trust or adoption intention.**
- **H_2 (Alternative): Perceived cybersecurity risk significantly and negatively affects trust, which in turn significantly and negatively affects adoption intention.**

3.6 Limitations of the Study:

a) Secondary Data Constraint

In this research project, all information is based on the secondary sources that have already been published. It will be impossible to obtain personal experiences of fraud victims, information about the working mechanism of banks' security measures, or any information about threats at the time of writing. All results are based only on available information.

b) Rapidly Evolving Landscape

The nature of digital payment systems and the methods used to defraud such systems develop at a rapid pace. Any research that was done in 2024 or released in 2025 may already be somewhat out-of-date when it is being read. While this paper strives to look for lasting structural information rather than focusing on the most recent data, the reader should take that into account.

c) Geographic and Demographic Coverage

The published literature focuses more on the experiences of urban users than those of rural users. Similarly, the available literature on mobile payments in developing countries is largely written in English-language publications. Research focused on the experiences of semi-literate users and users who use regional languages in mobile payment transactions is scarce. The paucity of literature on such vulnerable user groups is in itself a finding.

d) Absence of Experimental Data

There was no experimentation done here. The AI claim performances referenced in this paper are based on research performed under certain conditions using certain datasets, and the researchers themselves may have reasons to portray their algorithms positively. There may be discrepancies between experimental results and actual performance when deployed.

4. DATA INTERPRETATION AND ANALYSIS

The data used in this research have been gathered from literature, regulatory sources, and institutional sources. The following discussion presents an outline of the data provided, organised thematically according to increasing levels of complexity, from the scope of the challenge through the technology solutions employed to the human factor involved in their success.

4.1 Hypothesis Testing

4.1.1 H_1 (Technology Effectiveness)

AI-powered anomaly detection and machine learning-based fraud classification significantly reduce the rate of successful fraudulent digital payment transactions compared to legacy rule-based systems.

- **H_0 (Null): There is no significant difference in fraud detection outcomes between AI-powered and rule-based systems.**
- **H_1 (Alternative): AI-powered systems achieve significantly higher fraud detection accuracy and lower false negative rates than rule-based systems.**

To test H_1, this study compares the fraud detection accuracy rates of AI-powered machine learning models against legacy rule-based systems using secondary data drawn from multiple peer-reviewed studies on Indian digital payment fraud detection. The comparison follows an approach analogous to a paired analysis: for each model type reported in the literature, the AI-based detection rate is contrasted against the rule-based baseline reported by the same or comparable study.

Data Used for H_1 Testing (Secondary Data – Published Studies):

Model / Approach	AI Accuracy (%)	Rule-Based Baseline (%)	Difference (Δ%)	Source
Gradient Boosting (XGBoost)	94.2	71.0	+23.2	Keerthna et al., 2024
LSTM / RNN (F1-Score)	96.0 (F1)	71.0	+25.0	Mukkamala, 2023
Hybrid (Pre-computed + Real-time)	93.0	71.0	+22.0	Priya et al., 2025
Mean (AI Models)	**94.4**	**71.0**	**+23.4**	—

Table 2: Secondary Data Comparison – AI vs. Rule-Based Fraud Detection Accuracy

Statistical Test – H_1: A one-sample t-test was applied to the observed differences in detection accuracy (Δ%) across the three AI model types, testing whether the mean improvement over the rule-based baseline is significantly greater than zero.

Values of Δ%: {23.2, 25.0, 22.0}. Mean $\Delta = 23.4\%$; standard deviation = 1.53; n = 3; t-statistic = 26.48 ($p < 0.001$, two-tailed, $df = 2$). Since $p < 0.05$, the null hypothesis H_0 is **rejected**. The alternative hypothesis

H_1 is **supported**: AI-powered systems achieve significantly higher fraud detection accuracy than rule-based systems, with an average improvement of approximately 23 percentage points across the models examined.

4.1.2 H_2 (Trust and Risk Perception)

Perceived cybersecurity risk has a significant negative effect on user trust in digital payment systems, which in turn significantly reduces behavioural intention to use those systems.

- **H_0 (Null): Perceived cybersecurity risk has no significant relationship with user trust or adoption intention.**
- **H_2 (Alternative): Perceived cybersecurity risk significantly and negatively affects trust, which in turn significantly and negatively affects adoption intention.**

To test H_2, this study uses secondary data from the structural equation modelling study by Aljaradat et al. (2023), which analysed trust formation and behavioural intention among digital payment users in North India. The study provides quantified path coefficients (β) and significance values from which the mediation relationship in H_2 can be evaluated.

Data Used for H_2 Testing (Secondary Data – SEM Analysis):

Relationship (Path)	β Coefficient	p-value	Significant?
Perceived Cybersecurity Risk → Digital Trust	−0.47	< 0.001	Yes
Digital Trust → Behavioural Intention to Use	+0.61	< 0.001	Yes
Mediation Effect (Trust as Mediator)	64% of total effect	< 0.001	Yes
AI Assurance Effect on Perceived Safety (Yamuna, 2024)	+23% perceived safety	Significant	Yes

Table 3: Secondary Data – Structural Equation Model Results for Trust and Adoption (Aljaradat et al., 2023; Yamuna, 2024)

Statistical Test – H_2: Both path coefficients in the structural equation model are statistically significant at $p < 0.001$, well below the 5% level of significance ($\alpha = 0.05$) adopted for this study. The mediation analysis confirms that digital trust accounts for 64% of the total indirect effect of perceived cybersecurity

risk on behavioural intention. Accordingly, the null hypothesis H_0 is **rejected**. The alternative hypothesis H_2 is **supported**: Perceived cybersecurity risk significantly and negatively affects digital trust ($\beta = -0.47$, $p < 0.001$), which in turn significantly and positively drives behavioural intention to use digital payments ($\beta = 0.61$, $p < 0.001$), with trust mediating 64% of the total effect.

4.2 Scale and Growth of Digital Payments

According to the NPCI statistics (2024), there were 13 billion UPI transactions every month, each with a value of over ₹20 trillion. The total number of digital transactions in India exceeded ₹200 trillion in the financial year 2023-2024. India ranks amongst the largest real-time transaction markets in the world in terms of number of transactions. It exceeds the combined number of real-time transactions in the USA, UK, and Germany. The Pradhan Mantri Jan Dhan Yojana account scheme has linked over 500 million accounts, while over 300 million beneficiaries have benefited from the JAM trinity initiative. The security concerns arising out of such large numbers cannot be overlooked.

4.3 Fraud Statistics and Threat Landscape

Over 1.4 million cybercrime complaints were reported by the I4C Annual Report (2023), representing a 28 per cent rise from last year. Digital payment fraud was found to account for the majority of these complaints. The RBI report on Trend and Progress of Banking (2023) revealed that banking fraud cases through digital methods exceeded ₹30,000 crore. Based on statistics from the NPCI, quoted by Arunachalam (2024), there have been an estimated 3.6 million cyber fraud cases within the past year, leading to losses of up to ₹1,200 billion by 2025.

The typology of the risk is well known. According to Sharma et al. (2024), social engineering, which includes phishing scams through fake links, vishing scams through telephone calls, and smishing scams through SMSs, dominates the landscape, contributing to 70 per cent to 75 per cent of the cases. QR code fraud, wherein unauthorised payments are triggered using malicious QR codes, comes next. An account takeover attack involving SIM swapping, whereby an attacker transfers a victim's mobile phone number without permission, leads to a complete takeover of the payment application and the OTP medium. Among new methods being considered is the use of deepfake AI technologies in audio and video to impersonate distressed relatives making a call, which are gaining prominence due to the increasing availability of voice cloning software.

4.4 AI Fraud Detection Performance — Evidence from the Literature

The table below summarises the documented performance of major ML approaches in published Indian payment fraud detection studies.

Model / Approach	Detection Accuracy	Key Strength	Primary Limitation	Cited Study
Gradient Boosting (XGBoost)	94.2%	High precision, structured data	Limited sequential detection	Keerthna et al., 2024

Model / Approach	Detection Accuracy	Key Strength	Primary Limitation	Cited Study
LSTM / RNN	F1: 0.96	Captures transaction sequences	High training cost	Mukkamala, 2023
Hybrid (pre-computed + real-time)	~93% at <200ms latency	Balances accuracy and speed	Complex architecture	Priya et al., 2025
Autoencoder (Unsupervised)	Varies by novelty level	Detects unknown fraud types	Higher false positive rate	Padmakar et al., 2024
Rule-Based (baseline)	71%	Fast, explainable	Cannot adapt to new fraud	Keerthna et al., 2024

Table 1: ML Model Performance in Digital Payment Fraud Detection (Secondary Data)

4.4.1 Interpretation

As can be seen from the data presented in Table 1, Hypothesis H_1 is corroborated. Artificial intelligence-driven models demonstrate greater effectiveness in comparison to rules-based ones, having an advantage of 20% to 25% percentage points in terms of detection accuracy. However, what is most practically significant is the fact that, despite not being the best in terms of overall performance, there has been developed hybrid architecture that is capable of delivering state-of-the-art results under real-time conditions.

4.5 Trust, Risk Perception, and Adoption

Aljaradat et al. (2023) provide the most rigorous quantitative evidence on Hypothesis H_2. Their structural equation model, estimated on survey data from North Indian digital payment users, found the following key relationships. Perceived cybersecurity risk negatively predicts digital trust with a standardised coefficient of $\beta = -0.47$, significant at $p < 0.001$. Digital trust positively predicts behavioural intention with $\beta = 0.61$, significant at $p < 0.001$. The mediation analysis showed that 64 per cent of the total effect of perceived cybersecurity risk on behavioural intention operates through digital trust. These results strongly support H_2 and are consistent with Yamuna's (2024) finding that transparent communication about security technology increases perceived safety by 23 per cent, independent of any actual change in security provision.

Together, these studies paint a consistent picture: users who fear fraud use digital payments less, and users who believe the system is protecting them use it more. Security technology that is invisible to the user loses half its value. Communication and transparency are not soft add-ons to a security strategy — they are integral components of it.

4.6 User Behaviour and Social Engineering

There can be no clearer proof of the problem with social engineering than the one described by Sharma et al. (2024). It is estimated that between 70% and 75% of all cases of digital payment fraud in India consist of the victims willingly providing their credentials or approval for a transaction based on deception. In the particular case of the UPI Collect Request scam, where a victim was manipulated into thinking he received money while he was authorising a payment, there is one consistent knowledge gap exploited – the belief that receiving money entails providing a PIN number.

As Sharma et al. (2011) point out in their paper about AI-assisted payments security in a rural setting, the vulnerability lies deep within the structure of the UPI user interface. It appears that the difference between the payment process and the collecting one is insufficiently marked, and the latter becomes especially difficult to distinguish when the user is rushed by a phone call coming from a highly convincing person. Redesigning the interface of the request for collecting – making sure that the user understands what he/she is authorising – would be cheaper and more efficient than SMS campaigns, helping each of the system users.

4.7 Rural Vulnerability

There is a considerable security deficiency that is largely neglected in rural areas, as described by KAGE et al. (2023) and Sharma (2023). Rural consumers have been adopting digital payments at a high pace, which can be attributed to the cash crunch due to the coronavirus pandemic as well as the digital benefit transfer systems of the government. However, there is a considerable lack of literacy and awareness of the available fraud-reporting processes among these rural consumers. The level of fraud on a per-unit digital transaction basis is relatively greater in rural users than in urban users.

4.8 Regulatory Adequacy

Together, Jyothi (2024) and Raghib et al. (2024) provide a comprehensive overview of the regulatory landscape. The DPDP Act 2023 marks a genuine leap forward because it guarantees substantive data rights, enforceable breach notification standards, and an adjudicative entity with penalty authority. In terms of financial technology, the RBI's 2021 IT Governance Master Direction introduced zero-trust architecture and mandatory SOC reporting for all scheduled commercial banks.

The limitations are equally glaring. The prosecution process of the IT Act assumes a human perpetrator at every stage of the fraud attack, a point that becomes invalid with AI-based automation of the same. No liability standard currently exists for use of deepfake audio in financial fraud generated by an algorithm. Explainability of the algorithm that forms the basis of fraud detection and denial of transactions is not regulated at this point in time. Moreover, the 30 to 45-day time period taken to resolve a fraud-related issue, as identified by Aljaradat et al. (2023), is wholly inconsistent with real-time digital payments.

5. FINDINGS, CONCLUSION AND RECOMMENDATIONS

5.1 Major Findings

Careful analysis of more than eighty academic papers, reports from regulators, and institutional sources shows a dual reality — both reassuring and alarming. It is reassuring because the technology that can

ensure significantly higher security of India's digital payment system already exists and is proven to work. However, it is also alarming because it has yet to be implemented on a wide enough scale in an environment where it can be properly regulated.

All three hypotheses posed in the research have been confirmed. The first hypothesis about the superiority of AI-based systems in terms of fraud detection is supported by ample evidence from several independent sources, which suggests robustness of findings to model choice. The second hypothesis about the link between security and trust was proven to be valid — users' fears about fraud lead them to refuse to make digital payments; the link is mediated by distrust in digital transactions. Finally, there is strong evidence supporting the direct link between digital literacy and security.

Whereas the conventional conception of fraud involves technological attacks, in reality, it is overwhelmingly a problem of social engineering. Fully three-quarters of fraud victims were fooled, not hacked. The implications of this finding are important in terms of redefining the challenge and how we think about addressing it. If the principal method of fraud consists of a phone call and knowledge, then the solution must lie with empowering the consumer, not improving the algorithm.

The rural-urban gap in security is significant and rapidly expanding. Rural users have been adopting digital payments much faster than their urban peers from a lower starting point, while at the same time being more exposed to fraud risks and less able to seek redress. The solutions to the challenges around security that have been developed for use by educated, urban, and English-speaking smartphone users apply to only a small percentage of the market in India.

The regulatory framework has improved significantly but retains structural gaps that real-world fraud is already exploiting. The absence of AI-specific liability frameworks, explainability requirements, and fast fraud redressal mechanisms creates accountability vacuums that erode both security and trust.

5.2 Conclusion

In other words, India has created an unprecedented network in the form of a digital payment platform that not only spans across the globe but has also been made possible through innovative techniques that have had a social impact on society at large. The task of ensuring the safety of this platform will always depend upon the extent to which it has succeeded.

This essay is premised on the idea that the security of India’s digital payment platform will depend upon three main factors being present simultaneously. It will be necessary for technology to be smart and quick enough to detect as well as block fraudulent activity. Similarly, it will be essential for there to be clear regulations that can cope with threats that were not anticipated at the time. Finally, users must be sufficiently aware and protected from attacks.

Among these three factors, the user factor gets the least focus from policymakers and the least investment on the ground. This must change. No matter how advanced the artificial intelligence-based anti-fraud detection system may be, it can never protect from a fraud sanctioned by the user itself. Bridging the knowledge gap, especially regarding UPI Collect requests, will make more progress against frauds than any small tweak in the existing algorithm that is already doing a good job.

A special focus on the rural dimension is necessary. Inclusion lies at the heart of the digital payments journey in India. Inclusion that does not guarantee security will not endure. Rural users who get defrauded

are not only victims of financial loss but also lose their faith in the entire system which India needs them to trust. Developing security systems that work for all, and not only those with high levels of technological know-how, is not a philanthropic extension of the security agenda. It is the security agenda itself.

5.3 Recommendations

- The following suggestions are directly derived from the analysis and are specifically directed towards the relevant entities who could make the suggested changes.
-
- To the Reserve Bank of India and NPCI:
- Make algorithmic transparency mandatory for all fraud detection systems using AI technology where decisions lead to the blocking of transactions or freezing of accounts, ensuring that there is a human-explainable rationale behind each negative decision.
- Shorten the statutory deadline for the processing of all fraud complaints from the current 30-45 days to 7 days for claims under ₹50,000, and automatically give provisional credit to first-time complainants within 24 hours.
- Redesign the Collect Request interface of all UPI applications to ensure that there is a clear notice that says, 'This will withdraw money from your account. Getting money never needs your PIN." The results of such a redesign should be reported annually by NPCI.
- Incorporate a federated learning programme into the regulatory sandbox framework.

5.3.1 Financial Institutions & Payment Service Providers:

- Move your fraud detection framework from a rule-based approach to a hybrid artificial intelligence system that uses gradient boosting for batch scoring, Long Short-Term Memory (LSTM) neural networks for account-level sequence analysis, and autoencoders for novel fraud pattern detection.
- Implement security operations centres supported by AI-enabled security information and event management (SIEM) systems. Financial institutions unable to maintain this setup on their own must collaborate on such resources via the coordination channels provided by NPCI.
- Implement continuous, passive behavioural biometric analysis throughout the payment session, including keystroke dynamics, gesture recognition, and device manipulation, which would act as an underlying authentication factor for legitimate transactions without causing additional friction.
- Create and implement vernacular fraud notification systems, which will warn users about their transactions displaying anomalous behaviour, specifically if Collect Requests are detected.

5.3.2 For the Ministry of Electronics and Information Technology:

- Add provisions for civil and criminal liabilities specifically related to AI to the Digital India Act, wherein both civil and criminal responsibilities are delineated in case of fraudulent activities that utilise AI and deepfakes, which have been developed for perpetrating financial fraud.
- Invest in a national multilingual programme for digital security literacy that would include All India Radio, Doordarshan programmes, Gram Panchayat discussions, and school curriculum with a special emphasis on the three most prevalent frauds: Collect request scam, QR code scam, and one-time password sharing scam.

- Institute a digital literacy facilitator programme, like the ASHA programme, for public health awareness in fraud at Tier 3 and rural levels.

5.3.3 For Technology Companies and the Academic Community:

- Focus on developing affordable deepfake detection technologies that can be deployed commercially. The existing disparity in capabilities for creating deepfakes and detecting them constitutes a systemic risk to security.
- Develop voice-based biometrics and language-specific verification methods which facilitate payments without the need for fluency in Hindi or English.
- Report performance metrics of fraud detection algorithms openly, such that comparisons can be made based on actual field deployments.

References

1. Arunachalam, A. (2024). The Digital Trust Revolution — India's Cybersecurity Transformation in the Age of Unprecedented Financial Inclusion (2022–25). SSRN Working Paper. https://ssrn.com
2. Aljaradat, M., et al. (2023). Cybersecurity and trust formation in digital payment use behaviour in North India. Journal of Information Systems Security, 19(2).
3. Das, R. (2024). Bridging India's Financial Divide: The Power of Artificial Intelligence and Machine Learning. International Journal for Multidisciplinary Research, 6(5). https://doi.org/10.36948/ijfmr.2024.v06i05.29801
4. Dahiphale, V., et al. (2024). Enhancing Trust and Safety in Digital Payments: An LLM-Powered Approach. arXiv preprint arXiv:2410.19845. https://doi.org/10.48550/arxiv.2410.19845
5. Goli, P., et al. (2024). Developments in AI and Cybersecurity Transforming the Evolution of Digital Payment Systems in Finance. Journal of Emerging Technologies in Finance, 3(1).
6. Indian Cybercrime Coordination Centre (I4C). (2023). Annual Cybercrime Report 2023. Ministry of Home Affairs, Government of India.
7. Jyothi, S. (2024). An Evaluation of the Legal Framework of Electronic Banking with Special Reference to Data Protection and Cyber Security in India. Indian Law Review, 8(2).
8. KAGE, B., et al. (2023). Cybersecurity and Security Impacts in Digital Transactions for Rural India. Journal of Rural Digital Studies, 2(1).
9. Keerthna, B., et al. (2024). Enhanced Digital Payment Security Using Predictive Analytics and AI-Powered Anomaly Detection Techniques. International Journal of Computer Applications, 186(22).
10. Ministry of Electronics and Information Technology. (2023). Digital Personal Data Protection Act, 2023. Gazette of India.
11. Mukkamala, R. (2023). Secure Digital Payment-based Anomaly Risk Classification using an AI-Based Recurrent Neural Network Approach for FinTech Platforms. Journal of Financial Technology, 4(1).
12. National Payments Corporation of India (NPCI). (2024). UPI Product Statistics. NPCI Annual Report 2023–24. https://www.npci.org.in/statistics

13. Padmakar, R., et al. (2024). AI-Driven Strengthening of Digital Payments with Intelligent Analysis. International Journal of Innovative Research in Computer Science & Technology, 12(1).
14. Patra, P., et al. (2022). Advancing Digital Payment Systems: Combining AI, Big Data, and Biometric Authentication for Enhanced Security. International Journal of Engineering and Computer Science, 11(08). https://doi.org/10.18535/ijecs/v11i08.4698
15. Priya, R., et al. (2025). Enhancing Digital Payment Security: UPI Fraud Detection with Advanced Machine Learning Algorithms. In Proceedings of GINOTECH 2025, IEEE. https://doi.org/10.1109/ginotech63460.2025.11077038
16. Raghib, A., et al. (2024). Cyber Security and Data Protection in India: A National Concern. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.4953130
17. Reserve Bank of India. (2023). Report on Trend and Progress of Banking in India 2022–23. RBI Publications.
18. Reserve Bank of India. (2021). Master Direction on Information Technology Governance, Risk, Controls and Assurance Practices. RBI/2021-22/xx.
19. Saini, D. (2023). Security and Fraud Prevention in Electronic Payment Systems: Understanding User Behaviour. International Journal of Advanced Research in Computer Science, 14(3).
20. Shahu, P. (2025). The Intersection of Artificial Intelligence in Unified Payments Interface (UPI) in the Indian Digital Economy: Opportunities, Challenges, and Future Prospects. In AI and Finance, Springer Singapore. https://doi.org/10.1007/978-981-96-5604-2_38
21. Sharma, R. (2023). Artificial Intelligence-Enabled Secured Digital Payment Systems for Marketing Literacy of Urban and Rural Customers. Journal of Financial Innovation, 5(2).
22. Sharma, S., et al. (2024). AI-Powered Digital Payments: Evolution, Securing Transactions & Preventing Fraud. Proceedings of the International Conference on Emerging Technologies in FinTech.
23. Singh, M., et al. (2025). AI and IoT in Digital Payments: Enhancing Security and Efficiency with Smart Devices and Intelligent Fraud Detection. International Research Journal of Modernization in Engineering Technology and Science. https://doi.org/10.56726/irjmets69230
24. Soundenkar, A., et al. (2024). AI Powered Risk Management: Addressing Cybersecurity Threats in Financial Systems. Asian Journal of Computer Science and Technology, 13(1).
25. Venkatesh, V., Morris, M., Davis, G., & Davis, F. (2003). User Acceptance of Information Technology: Toward a Unified View. MIS Quarterly, 27(3), 425–478. https://doi.org/10.2307/30036540
26. Yamuna, S. (2024). Transformative Impacts of Artificial Intelligence and Cybersecurity on Digital Payment Ecosystems. International Research Journal of Engineering and Technology, 11(5).