International Journal of
Science, Architecture, Technology and Environment

Effectiveness of AI based fraud detection in UPI transaction

Anagha V K[1*], Jagannadhan S[2]

[1,2]*Post Graduate Diploma in Management (PGDM) Dayananda Sagar Business School*

[*]*Corresponding author- anaghanair090@gmail.com, jagansiji004@gmail.com*

DOI: https://doi.org/10.63680/ijsate0226059.051

Abstract

Despite its swift growth, India's Unified Payments Interface now faces rising threats even as it reshapes daily money transfers between people and companies. Instead of slowing down, scams like fake messages, stolen identities, altered payment codes, and manipulation through deception have grown alongside adoption rates. A closer look at artificial intelligence appears in this paper - named "Effectiveness of AI-Based Fraud Detection in UPI Transactions" - to assess whether smart systems can catch suspicious activity faster than older methods do. From spotting odd patterns to cutting down on lost funds, machine learning may quietly help users feel more confident when paying online. A closer look at how people view AI-driven fraud prevention begins with gathering answers from 150 individuals using UPI services. Because responses came via a fixed-format survey, insights emerge on awareness levels alongside feelings of safety and reliance on automated safeguards. Information beyond firsthand replies comes into play too - documents issued by the Reserve Bank of India, materials from NPCI, scholarly articles, and investigations within fintech shape context around rising digital scams and machine learning use in banking defense. When it comes to making sense of numbers, methods like proportional breakdowns, visual graphs, and side-by-side evaluations highlight tendencies tied to scam encounters, instant warning signals, precision in identifying threats, incorrect flags, and faith in system performance. Patterns take form not just through volume but through contrast and repetition across sources.Most people who took part in the survey view AI-powered fraud detection tools as more capable than older rule-driven methods. Real-time tracking, automatic warnings, and quicker spotting of unusual behavior helped boost trust among users. Yet doubts linger - some worry about personal information being exposed, legitimate payments getting wrongly stopped, or simply not knowing how these smart systems actually work. Even though artificial intelligence clearly strengthens UPI safety layers, lasting impact will rely on clearer system logic, fewer incorrect flags, better public understanding, and stronger safeguards for private data. These outcomes matter greatly for financial institutions, tech firms, oversight bodies, and scholars aiming to shape a reliable digital money landscape across India.

Keywords: Artificial Intelligence (AI), UPI Fraud Detection, Digital Payment Security, Machine Learning in Finance, Real-Time Transaction Monitoring

1. Introduction

The Unified Payments Interface (UPI) has fundamentally transformed India's financial landscape, evolving from a novel convenience into the primary pulse of the nation's economy. Right now, money moves differently across India. What began as just another way to pay has turned into how most people handle cash every day. By 2026, the number of these digital payments hits levels nobody expected. With so much activity happening fast, criminals see fresh chances to strike. Speed wins when moving money - yet that same speed helps scammers pull off theft before anyone notices. Old-style tools built to catch fraud start failing more often. They depend on fixed steps written by analysts long ago. But today's hackers shift methods quickly, using clever tricks instead of brute force. Their attacks feel personal, timed perfectly, slipping past defenses meant for simpler times.

Out of nowhere, tech changes how safety works - threats vanish before they grow. Moving forward, money handlers rely on tools tracking actions: where people log in, what devices they use, even how fingers slide across glass. Tiny details most ignore start linking directly to each purchase made. Behaviors seeming normal get lifted into view by unblinking online eyes. Many walk right by these quiet mismatches, never pausing once. Old programs freeze up right here. Out of nowhere, movement kicks in - no waiting. While you're still thinking, it's already reading your face. Occasionally pauses, quietly probing beneath the surface. All of it happens behind the scenes, complete long before money moves or everything just freezes.

When scams change, the system changes too - learning from each attempt, good or bad, feeding them into updated models without pause. Because it adapts on the fly, protection grows sharper while still letting real users move quickly. Mistakes drop. Legitimate actions stop getting flagged by accident. With every attack, insight forms. Quietly, behind the scenes, decisions refine themselves. Speed stays intact. Safety deepens. Fraud finds fewer openings. The framework holding up fast payments leans heavily on this quiet evolution. Not loud, never visible, but always adjusting. What keeps digital money transfers alive amid rising threats isn't armor made of steel, but one built from smart repetition. Each cycle makes the net tighter.

Even with progress, artificial intelligence still struggles in real-world situations today. By 2026, criminals began using smart algorithms to test bank safeguards, pushing defenders toward clearer decision-making processes. Because of this shift, systems now must show how they reach conclusions, not just act on hidden logic. One big issue? Alerts triggered by normal behaviour - like a grandparent sending money to family - that seem suspicious but aren't. Instead of relying on isolated data, banks now learn from each other through decentralized methods that protect personal details. Through shared insights built across institutions, the payment network strengthens without exposing individual records. Trust holds when speed meets safety, especially as digital payments vanish into everyday life unnoticed.

Review of Literature

1. Research Team, International Journal of Innovative Science and Research Technology (IJISRT), 2025. Review: This study proposes a hybrid fraud detection system combining rule-based approaches with machine learning. It highlights the effectiveness of anomaly detection and behavioural biometrics in monitoring user interactions to distinguish between fraudulent and genuine UPI activity in real-time.

2. Sharma et al., International Journal of Creative Research Thoughts (IJCRT), 2024. Review: The authors developed an adaptive weighted fusion classifier integrating Random Forest, Naive Bayes, and SVM. The study demonstrates that this ensemble approach minimizes false positives and effectively adapts to

emerging fraud patterns in UPI datasets

3. Research Team, Journal of Emerging Technologies and Innovative Research (JETIR), 2025. Review: This paper evaluates the performance of Random Forest and XGBoost algorithms. It concludes that Random Forest achieved the highest accuracy (90.75%) in categorizing transactions based on amount, frequency, and user behaviour.
4. Research Team, International Journal of Engineering Research & Technology (IJERT), 2025. Review: This survey provides a taxonomy of fraud detection methods, comparing ML models like XGBoost and SVM. It identifies the lack of real-world labelled datasets as a primary challenge and suggests ensemble stacking as a strong potential solution.

5. IEEE Conference Authors, IEEE Xplore, 2025. Review: Focusing on QR code tampering and transaction fraud, this research utilizes Random Forest and Isolation Forest algorithms. It emphasizes the need for AI to detect unauthorized changes in QR codes which are a common entry point for UPI fraud.
6. Vijay Bhaskar et al., International Journal of Science, Engineering and Technology (IJSET), 2026. Review: A very recent study that utilizes four popular algorithms (Random Forest, Logistic Regression, Decision Tree, and SVM). It finds that continuous learning from new transaction data allows the system to improve detection accuracy over time.

7. Research Team, International Research Journal of Engineering and Technology (IRJET), 2025. Review: This paper proposes a hybrid CNN-SVM architecture. The CNN component extracts complex features from transaction data, while the SVM classifier distinguishes between legitimate and fraudulent transactions, suitable for high-throughput real-time monitoring.

8. Bhramanand Sethi et al., Atlantis Press, 2025. Review: The study focuses on a Random Forest model trained on transaction details like sender activity and device data. It reports a high accuracy of 96% and emphasizes the importance of data preprocessing (handling missing values, scaling) for model performance.

9. Gupta et al., ResearchGate, 2025. Review: Deviating from traditional ML, this research uses a BiLSTM (Bidirectional Long Short-Term Memory) and Transformer-based encoder. The model captures temporal dependencies in transaction sequences, achieving near-perfect accuracy (99.90%) on a large dataset.

10. Research Team, ResearchGate, 2025. Review: This paper evaluates multiple classifiers but proposes a CNN framework as superior for processing large-scale transaction data. It highlights the capability of deep learning to automate feature extraction, reducing the need for manual feature engineering.

11. Research Team, International Journal of Innovative Research in Technology (IJIRT), 2025. Review: This unique study integrates Natural Language Processing (BERT) and Graph Neural Networks (GNN) with crowdsourced user reports. It aims to detect scam networks and flag suspicious UPI IDs before a transaction is even initiated.

12. Research Team, International Journal for Research Trends and Innovation (IJRTI), 2025. Review: The authors apply the Random Forest algorithm to historical datasets. A key contribution is the analysis of "user interaction" where the system provides interactive analytics to users, explaining why a

transaction might be flagged as risky.

13. Research Team, International Journal of Engineering Technology and Management Sciences (IJETMS), 2025. Review: This review compares ANN, LSTM, and GRU models. It argues that while ANNs are foundational, LSTM and GRU are superior for fraud detection because they can analyse sequential patterns and time-series data effectively.

14. Research Team, International Journal for Research in Applied Science and Engineering Technology (IJRASET), 2024. Review: The paper reviews three categories of solutions: synthesis-based, algorithmic, and data-layer. It discusses the challenge of imbalanced datasets and proposes class-weight-based approaches to improve the identification of rare fraudulent events.

15. Research Team, International Journal of Research and Analytical Reviews (IJRAR), 2025. Review: Focusing on rural adoption of UPI, this paper uses KNN and Decision Trees to classify transactions based on bank account names and amounts. It emphasizes the need for simple yet effective models that can run efficiently to protect less digitally literate users.

Research Gap

- **The Explainability Crisis**
  Most high-performing deep learning models act as "black boxes," making it difficult for regulators and users to understand the specific logic behind why a legitimate transaction was blocked.
- **Adversarial AI Attacks**
  There is a significant research gap in defending against "Fraudster AI," where criminals use their own machine learning models to probe and find loopholes in a bank's defensive algorithms.
- **DigLatency vs. Accuracy Trade-off**
  While AI models are highly accurate, maintaining this precision within the mandatory sub-500 millisecond response time required for a seamless UPI experience remains a technical bottleneck.
- **Limited Cross-Platform Intelligence**
  Fraud detection is often siloed within individual banking apps, leaving a gap in identifying "mule accounts" that move stolen money rapidly across different payment providers.
- **Detection of Social Engineering**
  Current AI excels at catching technical hacks but struggles to intervene when a legitimate user is manipulated into voluntarily authorizing a fraudulent transaction through psychological pressure.

Objectives of the study

1. To examine the effectiveness of artificial intelligence in detecting and preventing fraudulent transactions within the UPI ecosystem.
2. To assess the impact of AI-based fraud detection systems on user trust, confidence, and satisfaction in using UPI for digital payments.
3. To compare AI-driven fraud detection mechanisms with traditional rule-based systems in terms of accuracy, speed, and reduction of false positives.

Statement of the Problem

Digital payments have spread fast through India, thanks to how quickly UPI usage has grown - bringing speed and ease to money transfers. Still, that surge opened doors for more online scams like phishing attacks, stolen identities, phony payment demands, along with sneaky account intrusions. Older anti-fraud tools built on fixed rules tend to fall short when facing clever, shifting tricks used by scammers these days. That weakness ends up costing customers, banks, even tech platforms real money in damages.

Even when machines scan money moves, spotting odd behavior fast does not always mean fewer mistakes happen. What counts is whether clever software tells true buys apart from fake ones without slowing things down. Still, doubts linger - over who gets access to your data, whether guidelines favor some, what drives decisions, plus how sharply new fixes spot hidden threats now. The heart of it stays clear: measuring how tightly clever systems protect online money moves weighs just as heavily as holding trust steady and transactions flowing.

Limitations of the Study

- Limited Access to Confidential Data: One reason could be tighter rules around user privacy. Accessing live UPI payments might not always be allowed. Security laws often block immediate data sharing. Researchers sometimes face delays because of these barriers. What shows up in reports may already be outdated. Real-time flow is hard to capture under current policies.
- Rapidly Changing Fraud Techniques: Fresh tricks pop up every time tech takes a leap forward. Fresh tricks in fraud can leave yesterday's smart systems struggling to keep up. Fraud patterns could shift beyond what the research reflects.
- Reliance on Information Collected by Others: One way this study could work is by using existing materials like published reports or journals instead of gathering new information. Sometimes older analyses - such as detailed examples from past situations - are enough to build understanding. Information already out there, including summaries written by others, might form the core of what's examined here.
- Limitations in Tech and Infrastructure: Good AI needs solid tech setup plus trustworthy information. When computers work well, results improve - especially if the details fed into them are clear and correct. Without reliable hardware, even smart programs fail quietly behind the scenes. Data that is messy or incomplete weakens every outcome slowly over time. Performance rises only when both pieces fit tightly together without gaps.
- Limited Geographic Scope: Perhaps the research zeroes in on UPI payments, say, within one particular country. Though it could just as easily center on a smaller regional level instead. Fraud trends shift under new skies - AI's reach changes shape across regions. Where tricks evolve, tools adapt differently. So, results might not fit every situation. Still, they reflect a specific context clearly.

Research Methodology

Type of Research

The study is descriptive in nature. It is descriptive because it describes the user perceptions and experiences by use of graphs.

Sampling Technique

One way this study gathered information was through whatever sources were easiest to reach. Instead of using random picks, it leaned on what was already out there about AI spotting fraud in UPI payments. Much of the material came from reports and examples that happened to be handy. Because access mattered more than balance, the process skipped strict selection rules. What showed up in the analysis often just depended on what could be found without much effort.

With help from established organizations and verified data, this research tries to show how well artificial intelligence spots fraud in UPI payments across the country. Instead of guesses, it leans on proof to measure real impact on financial scams. Clear results come from sources that have been checked before. What emerges is a picture based not on hopes but on what has actually happened. Through solid information, the work outlines where AI helps - and where it might fall short - when guarding digital money transfers.

Sample Size

This look into things doesn't pin itself to how many people respond, since it leans more on what's already been gathered instead of asking fresh questions. Year after year numbers from across the country shape the heart of the work - tracking where UPI scams popped up, money that went missing, how often those tricks were caught, along with spots where smart software started watching for threats.

Sample Design

This study builds on information gathered mostly from existing records, zeroing in on artificial intelligence used to catch fraud in UPI payments. Rather than picking at random, data was drawn deliberately - national figures on scams and AI results - from trusted bodies like the RBI and NPCI. Year after year, these outlets deliver numbers that hold up under scrutiny.

Actual collection of data

Questions built around real UPI usage helped learn how users spot risky transactions and whether they notice scams. People answered online, using secure links that fit into daily routines without hassle. Each part of the form explored feelings about instant warnings, AI spotting fraud, and trust in payment safeguards.

Tools Used

There were a variety of tools used both descriptive and inferential, to assess the outcomes of the study:

- Google Forms (for distributing and collecting the questionnaire digitally).
- Graphs and Charts - we used bar graphs, pie charts, and column graphs to display data visually to help with interpretation because that is an easier way to read data visually.

This study was able to include numerical accuracy using descriptive statistics along with the visual representation of data to help provide some clarity in interpreting the data.

Method of data analysis

The Data generated from 150 respondents, via Google Forms, and the secondary data was compiled into

Google Sheets and imported into an Excel file for deeper exploration. The spreadsheet/Excel file was cleaned removing duplicates and modified to recognize formatting issues and inconsistences. Using the various provided functions in Excel, such as filters, pivot tables, and conditional formatting, the responses were delivered into a specific category (frequency and type) and followed by further exploration. Some inclusions were associated, and charts - such as bar graphs and pie charts - were produced to help visualize trends and/or findings. This process helped ensure the findings were accurately interpreted and assisted in representing the findings through charts.

Limitations

- Limited Representation: A handful of people took part in the research, so their answers might not match what most UPI users across India face. Some groups notice scams faster than others, depending on how they use the system.
- Reduced Generalizability: A small group might not show what most people think. These outcomes could just mirror those who took part.
- Sampling Bias: When nearly everyone surveyed falls into the same age bracket, schooling level, or location, results might tilt one way. Because of that, judgments on how well AI spots fraud could miss the mark.
- Limited Range of Personal Experiences: One reason some patterns go unnoticed is a narrow group being studied. That can limit insight into how machines react when scams differ widely. Though rare cases appear, they still shape outcomes behind the scenes.
- Lower Statistical Reliability: Fewer data points tend to weaken how solid the number work turns out. Because of that, spotting clear patterns in whether AI catches UPI fraud well becomes harder

Inferences

- The survey shows that the majority (45.3%) of respondents are in the 21-30 years age group, indicating that UPI usage and awareness of AI fraud detection is highest among younger adults who are more tech-savvy.
- With 42.0% being Salaried Employee, the survey captures responses from working professionals who are regular UPI users, making their feedback on AI fraud detection particularly relevant.
- 40.0% of respondents use UPI daily, demonstrating high adoption and frequent usage patterns. This frequent use makes them well-positioned to evaluate fraud detection effectiveness.
- 38.0% of respondents are highly aware of AI-based fraud detection systems, showing good awareness levels among UPI users about security features protecting their transactions.
- The distribution shows 30.0% are not sure if they experienced fraud attempts. Combined with 'Yes, once' responses, this indicates moderate exposure to fraud attempts, highlighting the importance of robust detection systems.
- 38.0% find AI-based fraud detection effective, indicating strong positive perception of the technology's capability to identify and prevent fraudulent transactions.
- 32.0% receive real-time alerts sometimes, suggesting that while alert systems are functioning, there's room for improvement in consistency and frequency of notifications.
- 35.3% responded neutral when asked if AI fraud detection prevented financial loss. This suggests users may not have direct experience with prevented fraud, or are uncertain about the technology's role in protection.

- 40.0% are satisfied with UPI app security features, indicating overall positive sentiment toward current security measures including AI-based fraud detection.
- 38.0% agree that AI reduces false transaction blocks, suggesting the technology is effectively balancing security with user convenience by minimizing incorrect flagging of legitimate transactions.
- 40.0% mostly trust AI technology for securing transactions, demonstrating strong confidence in AI-based security measures among UPI users.
- A significant 42.0% strongly agree that AI systems are better than traditional methods, indicating strong user confidence in advanced AI approaches over rule-based detection systems.
- 38.0% agree that AI monitoring has improved their confidence in using UPI, showing the positive impact of AI fraud detection on user trust and adoption.
- An overwhelming 42.0% consider AI fraud detection very important for digital payments' future, emphasizing the critical role users expect AI to play in securing digital transactions.
- Overall, 38.0% rate AI fraud detection as excellent, reflecting strong satisfaction with current AI-based security measures in UPI transactions and validating the effectiveness of these systems

Findings

- Overall Insight:

Fraud detection using AI in UPI payments got good responses because it helps keep money safe, cuts down on scams, yet builds trust among people who use it. Still, for more users to rely on these tools, those offering UPI services must explain how things work, teach customers clearly, make sure systems run without glitches, while dealing quickly when alarms go off by mistake. As updates continue over time, artificial intelligence might just become a quiet backbone - helping protect India's growing world of digital cash transfers.
Recommendation

On the basis of the findings, the following recommendations are made:

- Improve User Awareness Initiatives: Not many realize what happens behind the scenes when a payment is flagged. Some confusion slips away once people see how smart systems spot odd patterns. When users learn what warnings mean, they tend to feel steadier using apps. Knowing small habits cut risks helps folks move easier online. Clear examples shared often make things click differently over time.
- Clearer Insights into Fraud Checks: A sudden notice appears when payments get stopped - telling users the reason helps them understand what happened. When alerts show up, a plain sentence about unusual behavior makes things clearer instead of leaving people guessing.
- Reduce False Positives: Every now and then, fresh adjustments keep AI systems sharp enough to let real transactions through without delay. When models learn better, people move through processes more smoothly because fewer valid attempts get flagged by mistake.
- Improved data privacy and security: Strong data encryption needs backing up with real attention to cybersecurity rules. Privacy policies explained well can ease worries users carry around their information being used wrongly. One wrong move here leaves trust cracked wide open.

Conclusions from the study

Fraud detection powered by artificial intelligence has made UPI payments notably safer, the research finds. Because tech keeps advancing, security can grow stronger alongside smarter tools and clearer guidance for users. When gaps are narrowed, banks and app developers help build lasting confidence in how people pay online. Support that responds well adds to a system were trust sticks around longer.

Direction for Future research

•Looking ahead, studies might explore how well artificial intelligence spots scams in UPI payments by pulling data from many areas, ages, because variety strengthens results. Wider information brings sharper clarity about who gets affected, why it happens. Instead of generalizing, comparing city dwellers with those in villages could reveal gaps in knowledge, levels of confidence, ways people run into online tricks.

•Maybe look into smart tech like deep learning later. Real-time risk tools might boost how well scams are caught. One study could compare old-school rules against new AI tricks at various banks. How those systems stack up matters. Watch how biometrics behave over time too. Some findings might show which method wins in speed or precision. Scammers change tactics often. Systems must shift just as fast. Long-term tests would reveal if artificial brains keep up. Results from different fintech spots add clarity. Not every model works the same everywhere. Performance shifts depending on where it runs. Learning happens slowly. So does improvement. What works today may fail tomorrow. That pattern repeats. Watching that cycle brings useful lessons. Each bank faces unique threats. Custom answers help more than one-size-fits-all fixes. Progress hides in details most miss. Patience pulls it forward.

•Looking ahead, one key path involves digging into ethics, how personal data is protected by rules, plus openness when machines make choices. What if researchers explored ways people learn about AI safeguards - how that shapes confidence in such tools? Thoughts on building stronger fraud defenses often come back to clearer policies and smarter awareness efforts. Safer online payments might depend less on tech alone, more on honest conversations about risk and control. Progress could hinge not just on code, but on whether users feel informed, respected, included.

Declaration of Conflicting Interests

The authors declare no potential conflicts of interest with respect to the research, authorship and publication of this article.

Funding

The author received no financial support for the research, authorship and publication of this article.Start new paragraph.
Do not change the format. Use this format for writing the manuscript.

References

- Reserve Bank of India. (2023). Annual Report 2022–23. Reserve Bank of India. https://www.rbi.org.in
- National Payments Corporation of India. (2023). UPI Product Statistics and Reports. NPCI. https://www.npci.org.in
- Arner, D. W., Barberis, J., & Buckley, R. P. (2016). The evolution of FinTech: A new post-crisis paradigm. Georgetown Journal of International Law, 47(4), 1271–1319.

- Ngai, E. W. T., Hu, Y., Wong, Y. H., Chen, Y., & Sun, X. (2011). The application of data mining techniques in financial fraud detection: A classification framework and an academic review. Decision Support Systems, 50(3), 559–569
- Bhatla, T. P., Prabhu, V., & Dua, A. (2003). Understanding credit card frauds. Cards Business Review, 1–15
- Ryman-Tubb, N., Krause, P., & Garn, W. (2018). How Artificial Intelligence and Machine Learning research impacts payment card fraud detection: A survey and industry benchmark. Engineering Applications of Artificial Intelligence, 76, 130–157.
- Kumar, A., & Gupta, S. (2022). Artificial intelligence in digital payment systems: Enhancing fraud detection and cybersecurity. International Journal of Financial Studies, 10(2), 45–60.