International Journal of Emerging Research in Computer Science
Volume 2 | Issue 2 | January 2026 | ISSN: 3139-101X (Online)
Available online at: https://ijercs.com/
DOI: https://doi.org/10.59828/ijercs.v2i1.7

AI-Based Fraud Detection in Digital Payment Systems

Dr. Sonali Sagar Gholve[1], **Bhagyashree Kishor Shinde**[2]

[1]Assistant Professor and Head, Department of Computer Science, Sarhad College of Arts, Commerce and Science, Pune

[2]Sarhad College of Arts, Commerce and Science, Pune

Received: 03 January 2026 | **Accepted:** 17 January 2026 | **Published:** 30 January 2026

Abstract

The rapid expansion of digital payment ecosystems, including mobile wallets, UPI transfers, internet banking, card-not-present transactions, and QR-code-based merchant payments, has significantly improved financial accessibility and transaction convenience. However, the same growth has also increased the attack surface for fraudsters, resulting in identity theft, phishing-enabled account takeover, synthetic identity abuse, mule-account activity, bot-driven transaction bursts, and adaptive social-engineering attacks. Traditional rule-based fraud detection systems remain useful for known patterns but struggle against evolving and low-latency fraud behaviors. This research paper presents a structured academic study of artificial intelligence (AI)-based fraud detection in digital payment systems, emphasizing machine learning, deep learning, anomaly detection, graph intelligence, and hybrid human-in-the-loop decisioning. The paper reviews existing literature, identifies major operational and technical challenges, proposes a layered AI detection framework, and synthesizes representative performance insights from prior studies and industry practice. The study concludes that robust fraud prevention in modern payment platforms requires a multi-model architecture that combines supervised risk scoring, unsupervised anomaly discovery, graph-based link analysis, explainable decision support, continuous feedback loops, and governance controls for privacy, fairness, and compliance.

Keywords: *AI, Fraud Detection, Digital Payment Systems, Machine Learning, UPI, Financial Technology, Anomaly Detection, Cybersecurity.*

1. Introduction

Digital payments have become a foundational component of the modern financial ecosystem. In India and many other emerging digital economies, the adoption of Unified Payments Interface (UPI), mobile wallets, QR-code merchant transactions, contactless card payments, and instant account-to-account transfers has changed the way consumers and businesses interact. The convenience of real-time settlement, frictionless onboarding, and 24×7 transaction capability has accelerated user trust and commercial dependency on digital financial infrastructure. At the same time, the concentration of high-volume, low-latency transactions has created a highly attractive target environment for organized cybercriminals and opportunistic fraud actors.

Fraud in digital payment systems is no longer limited to stolen cards or isolated phishing events. Modern fraud includes account takeover, device spoofing, SIM swap exploitation, mule-account orchestration, credential stuffing, remote-access manipulation, social engineering, merchant abuse, synthetic identity creation, and coordinated fraud rings operating across multiple channels. Attackers increasingly leverage automation and data leakage from external breaches, which allows them to

adapt quickly to static rule engines. As a result, financial institutions require detection systems that can identify subtle anomalies, temporal behavior shifts, and network relationships in near real time.

Artificial intelligence offers a significant improvement over purely manual and rule-based monitoring. Machine learning models can learn complex nonlinear patterns from historical transaction data, while deep learning approaches can capture sequential behavior, embeddings, and contextual interactions across multiple attributes. Unsupervised anomaly detection can identify previously unseen fraud, and graph analytics can reveal collusive relationships among devices, accounts, beneficiaries, IP addresses, merchants, and geolocations. Therefore, AI-based fraud detection is increasingly viewed as a strategic necessity rather than an optional innovation.

This paper investigates the role of AI in digital payment fraud detection from an academic and implementation perspective. The objectives are: (i) to explain the fraud landscape in digital payment systems; (ii) to review relevant research and practical model families; (iii) to propose a comprehensive AI-based detection architecture; (iv) to discuss evaluation metrics and deployment considerations; and (v) to identify limitations, governance requirements, and future research directions. The paper follows a structured journal-style format to align with standard academic submission expectations.

2. Literature Review

A substantial body of research has established that fraud detection in financial systems is fundamentally a class-imbalance problem, where fraudulent transactions represent a very small proportion of total payment volume. Bhattacharyya et al. (2011) demonstrated that cost-sensitive learning and sampling strategies are critical when fraud labels are sparse, because accuracy alone becomes misleading in highly imbalanced datasets. Their work influenced later studies that prioritized recall, precision, area under the ROC curve (AUC), and business cost over raw classification accuracy.

Dal Pozzolo et al. (2015) emphasized the importance of real-world evaluation conditions in payment fraud detection, especially the issue of delayed labels. In many banking environments, a transaction may be disputed days after authorization, meaning the training data available at time t is incomplete. This creates concept drift and label-latency problems. The authors argued that robust model monitoring, temporal validation, and adaptive retraining are necessary to sustain fraud model performance in production systems.

Jurgovsky et al. (2018) explored recurrent neural networks for credit card fraud detection and showed that sequence-aware models can capture behavioral patterns that are invisible to independent transaction classifiers. Their findings were important because payment fraud is often behavioral rather than static: a user's time-of-day, merchant type, device switching pattern, velocity bursts, and beneficiary novelty may collectively indicate compromise even when each feature appears normal in isolation.

Fiore et al. (2019) examined synthetic minority oversampling and feature engineering approaches in payment-card datasets, reinforcing the value of balancing techniques and domain-informed features such as velocity, transaction frequency, average ticket size deviation, merchant category transitions, and geographic dispersion. Their results suggested that strong feature design often matters as much as model complexity.

More recent literature has expanded beyond traditional supervised classification. Carcillo et al. (2021) described a practical industrial perspective in which streaming data pipelines, explainable AI, human investigator feedback, and hybrid architectures outperform isolated models. Their work highlighted the operational reality that fraud teams require not only scores, but also alert triage, explanation, feedback integration, and policy orchestration. Similarly, research on graph neural networks and graph-based link analysis has shown promise in detecting fraud rings and mule networks by modeling relationships among accounts, devices, and transaction paths rather than treating each event independently.

Studies on explainability, including work influenced by Lundberg and Lee (2017) through SHAP-based interpretability methods, have become increasingly relevant in regulated financial environments. Banks and payment service providers must justify declines, manual holds, and enhanced verification decisions, especially when customer experience is affected. Therefore, explainability is no longer an academic add-on; it is a practical requirement for compliance, auditability, and trust.

Overall, the literature indicates that no single model is sufficient for all fraud scenarios. Supervised gradient boosting remains highly effective for structured tabular scoring, sequence models improve account-behavior analysis, unsupervised models help discover novel fraud, and graph methods support ring detection. The strongest consensus in recent academic and

industry-aligned studies is that hybrid, layered, continuously monitored systems are the most effective design for digital payment fraud defense.

3. Problem Definition and Research Gap

Although many institutions deploy fraud controls, several gaps remain between academic models and production-grade digital payment systems. First, many published studies use static benchmark datasets that do not adequately reflect real-time streaming conditions, label delays, feature freshness constraints, or adversarial adaptation. Second, research papers often optimize for AUC or F1 score but under-emphasize operational metrics such as alert rate, false-positive cost, customer friction, manual review capacity, and loss-per-approved-transaction.

Third, a large proportion of earlier studies focus on card fraud, whereas modern digital payment ecosystems—especially instant transfer systems such as UPI and wallet rails—have different fraud dynamics. Fraud may be socially engineered, beneficiary-driven, device-linked, or session-contextual, which means contextual and behavioral features are more important than in traditional card-present settings. Fourth, the governance dimension is frequently underdeveloped: privacy, explainability, fairness, data minimization, model risk management, and regulatory defensibility are essential in actual deployments.

This paper addresses these gaps by presenting a broader digital-payment-specific fraud framework that combines machine learning, anomaly detection, graph analytics, and operational controls. Rather than treating the problem as only a binary classification exercise, the study frames fraud detection as a layered decision system that must balance security, speed, customer experience, and institutional accountability.

4. Research Methodology

This study follows a qualitative-analytical and design-oriented research methodology. It synthesizes academic literature, industry practices, and operational requirements to build a structured model of AI-based fraud detection in digital payment systems. The goal is not to claim proprietary experimental superiority on a private dataset, but to produce a rigorous, academically defensible framework that can be used for teaching, review, implementation planning, and further empirical research.

The methodology consists of five stages. In Stage 1, the fraud domain is characterized by identifying major digital payment channels and associated fraud typologies. In Stage 2, relevant AI techniques are mapped to those fraud types, including supervised classification, unsupervised anomaly detection, sequence learning, and graph-based analysis. In Stage 3, a proposed layered architecture is constructed to show how models can operate in real time and post-transaction monitoring. In Stage 4, performance evaluation metrics and deployment constraints are discussed. In Stage 5, practical limitations and future scope are derived from the synthesis.

The paper adopts an evidence-synthesis approach. Where numerical values are discussed, they are treated as representative patterns from literature and industry-aligned observations rather than a single proprietary benchmark. This is appropriate for a template-based academic paper when access to institution-specific payment data is restricted due to privacy, security, or legal constraints.

5. Fraud Typologies in Digital Payment Systems

A robust AI fraud detection framework begins with understanding fraud categories. In digital payment systems, fraud can be classified into several overlapping groups. Account takeover (ATO) occurs when attackers gain access to a legitimate user's account through phishing, credential stuffing, malware, SIM swap, or remote-access manipulation. Once control is established, the fraudster often adds a new beneficiary, changes device context, or initiates high-value or rapid-fire transfers.

Merchant and QR fraud occurs when attackers manipulate merchant onboarding, replace QR codes, redirect payments, or create fake merchant identities to receive funds. In peer-to-peer instant payment systems, beneficiary fraud and social-engineering scams are especially important. A customer may be induced into voluntarily authorizing a transaction under false pretenses, making the distinction between unauthorized and authorized push payment fraud operationally complex.

Mule-account and collusive fraud involve networks of accounts used to receive, layer, and disperse illicit funds. These schemes often evade simple threshold rules because each individual transaction may appear small or normal. Synthetic identity

fraud uses partially real and partially fabricated identity attributes to open accounts or pass weak onboarding checks. Device-sharing abuse, emulator-driven bot attacks, and scripted transaction bursts also appear in modern digital ecosystems. Each fraud type has different signals; therefore, feature engineering and model design must be aligned with the threat class rather than relying on a single generic detector.

6. AI Techniques for Fraud Detection

Supervised machine learning remains the core of most production fraud stacks. Logistic regression provides a strong baseline because it is fast, stable, and interpretable. Decision trees and random forests capture nonlinear interactions and are resilient to noisy data. Gradient boosting methods such as XGBoost, LightGBM, and CatBoost are often preferred in structured transaction data because they can model complex feature interactions, handle missingness effectively, and deliver strong ranking performance for risk prioritization.

Deep learning methods become valuable when behavior is sequential or multimodal. Recurrent neural networks (RNN), long short-term memory (LSTM), and gated recurrent unit (GRU) models can learn transaction sequences, temporal bursts, and repeated beneficiary patterns. Transformer-style sequence models are also gaining relevance in large-scale systems with rich event streams. Autoencoders and variational autoencoders can support anomaly detection by learning normal transaction representations and flagging reconstruction outliers.

Unsupervised and semi-supervised models are especially useful for emerging fraud patterns. Isolation Forest, Local Outlier Factor, clustering-based methods, and one-class classification can surface suspicious activity that lacks prior labels. In real environments, these models are valuable for analyst discovery, early warning, and shadow-mode monitoring even if they are not the final decline decision maker.

Graph-based techniques represent one of the most promising directions for advanced fraud defense. By constructing nodes for customers, devices, merchants, IP addresses, cards, bank accounts, beneficiaries, and sessions, and edges for interactions or fund flows, institutions can identify suspicious communities, circular money movement, device reuse across many identities, and abnormal connectivity patterns. Graph neural networks and link prediction models further improve the detection of organized fraud rings and mule structures.

Explainable AI methods such as SHAP, local surrogate explanations, rule extraction, and feature contribution analysis are essential companions to these models. In payment systems, an institution often needs to explain why a transaction was challenged, why a device was considered risky, or why an alert was prioritized. Therefore, the best technical design is usually a hybrid model ecosystem rather than a single algorithm.

7. Proposed AI-Based Fraud Detection Framework

This paper proposes a six-layer AI-based fraud detection framework for digital payment systems. Layer 1 is data ingestion and event normalization. It collects transaction events, user profile data, device telemetry, login sessions, beneficiary changes, IP metadata, merchant attributes, historical behavior aggregates, and external threat intelligence. A feature store or streaming enrichment service ensures that both real-time and historical features are available within strict latency budgets.

Layer 2 is rule-based and policy screening. Contrary to common perception, rules remain important. Sanctions checks, impossible travel heuristics, hard blacklists, velocity caps, suspicious device fingerprints, and regulatory blocks should be applied before or alongside model scoring. Rules are deterministic, auditable, and effective for known bad patterns. However, they should not be the only defense layer.

Layer 3 is supervised risk scoring. Here, gradient boosting models, calibrated probability estimators, and possibly stacked ensembles generate a fraud probability score using structured features such as amount deviation, time since last transaction, beneficiary novelty, device trust score, login friction, transaction velocity, merchant risk, and geospatial inconsistency. This layer handles the majority of known fraud patterns.

Layer 4 is anomaly and behavior intelligence. Unsupervised models examine deviations from normal user, device, merchant, and cohort behavior. Sequence models analyze whether the transaction is behaviorally plausible given recent account activity. For example, a new device followed by a password reset, beneficiary addition, and rapid high-value transfer within minutes is a strong behavioral risk pattern even if no single rule is triggered.

Layer 5 is graph intelligence and network risk. This layer scores relationships among accounts, devices, merchants, and fund flows to identify mule networks, fan-in/fan-out anomalies, device reuse across multiple identities, and coordinated fraud rings. The graph layer is particularly effective in post-onboarding abuse and organized financial crime scenarios.

Layer 6 is decision orchestration and feedback. Scores from prior layers are fused using business policies and calibrated thresholds. Low-risk transactions are approved silently. Medium-risk transactions may trigger step-up authentication, transaction delay, beneficiary cooling period, or customer confirmation. High-risk events may be blocked or routed to manual review. Investigator outcomes, customer disputes, and confirmed fraud labels feed back into the model lifecycle for retraining, threshold tuning, and drift monitoring. This layered design balances fraud prevention with customer experience and operational scalability.

8. Feature Engineering Strategy

Feature engineering is one of the most decisive factors in fraud model performance. Useful features can be grouped into transactional, behavioral, device, network, and contextual categories. Transactional features include amount, amount deviation from user median, merchant category, time since last payment, velocity over multiple windows, failed-attempt counts, and beneficiary novelty. Behavioral features include time-of-day consistency, day-of-week patterns, historical payment destinations, frequency of balance inquiry before transfer, and channel-switching behavior.

Device and session features include device fingerprint stability, OS version, app integrity signals, emulator detection, rooted-device indicators, SIM change recency, IP reputation, geolocation mismatch, and browser or SDK fingerprint continuity. Network features include number of accounts linked to a device, beneficiary fan-out, shared contact points, shared IP clusters, and fund-flow centrality. Contextual features may include holiday spikes, salary-day patterns, merchant campaign effects, or known fraud campaigns affecting specific regions or transaction types.

Feature freshness is also critical. In real-time payment systems, a stale feature may be worse than a missing one. Therefore, production design should distinguish between online features (available within milliseconds) and offline analytical features (available for batch monitoring). Strong governance of feature definitions, leakage prevention, and point-in-time correctness is necessary to avoid inflated offline results and weak live performance.

9. Model Evaluation Metrics

Fraud detection cannot be evaluated responsibly using accuracy alone because fraud is rare and false negatives can be costly while false positives can damage customer trust. Precision measures how many flagged transactions are actually fraudulent, while recall measures how many fraudulent transactions are successfully detected. F1 score balances both, but business teams often need more nuanced trade-offs than a single harmonic mean.

Area under the ROC curve (AUC-ROC) and area under the precision-recall curve (AUC-PR) are valuable ranking metrics, especially in imbalanced data. However, operational metrics are equally important. These include false-positive rate at a given approval target, alert volume per 100,000 transactions, manual review load, fraud loss per approved transaction, customer friction rate, step-up authentication success rate, and latency to decision. A model with slightly lower AUC may be preferable if it reduces unnecessary friction while preserving recall at high-risk thresholds.

Threshold selection should be business-aware rather than purely statistical. For example, a payment platform may use separate thresholds for silent approval, step-up verification, soft hold, and hard decline. Calibration is essential because a well-ranked model is not always a well-calibrated model. Probability calibration methods and periodic challenger testing help maintain stable decision quality over time.

10. Results and Discussion

Based on the literature synthesis and practical implementation patterns, several conclusions emerge. First, gradient-boosting models consistently perform strongly in structured payment datasets because they handle mixed feature types, nonlinear interactions, and missing values effectively. In many production environments, they remain the primary transaction-level scoring engine. However, their performance degrades when fraud patterns shift significantly unless retraining and drift management are applied.

Second, sequence-aware models provide substantial value in account takeover and behavioral fraud scenarios. A transaction may look legitimate in isolation but suspicious in sequence. For example, a password reset, device change, new beneficiary addition, and high-value transfer in a compressed time window may be highly indicative of compromise. Sequence models can capture this narrative more effectively than static models.

Third, unsupervised anomaly detection is best viewed as a complementary capability rather than a replacement for supervised models. It excels in surfacing novel or weakly labeled threats, but it can produce elevated false positives if used without context. Therefore, anomaly outputs should often be combined with rule filters, cohort baselines, or human analyst review.

Fourth, graph intelligence adds major value against organized abuse. Fraud rings often exploit the fact that single transactions appear benign while the network structure is abnormal. Device reuse across many accounts, shared beneficiary funnels, and circular fund movement become visible only when relationships are modeled explicitly. This is particularly relevant for mule-account ecosystems and merchant abuse.

Finally, the discussion confirms that the strongest fraud systems are not single-model systems. Effective deployments combine deterministic controls, real-time risk scoring, anomaly detection, graph intelligence, explainability, and feedback loops. The operational objective is not only to maximize model metrics, but to minimize fraud loss while protecting customer experience, maintaining regulatory defensibility, and scaling with transaction growth.

11. Tables and Figures

The following tables summarize the comparative characteristics of major AI techniques, representative performance patterns, and operational response strategies. In a final institutional submission, these tables may be complemented by actual charts, ROC curves, confusion matrices, or system diagrams based on proprietary or experimental datasets.

Table 1. Comparative Characteristics of AI Models for Digital Payment Fraud Detection

Model Type	Strengths	Limitations	Best Use Case
Logistic Regression	Fast, interpretable, stable probabilities	Limited nonlinear capture	Baseline real-time scoring
Random Forest	Robust to noise, nonlinear relationships	Heavier and less calibrated	General tabular fraud detection
Gradient Boosting	High performance on structured data	Needs careful tuning and monitoring	Primary production risk scoring
LSTM/GRU	Captures temporal behavior	Higher complexity, lower explainability	Account takeover and sequential abuse
Isolation Forest / Autoencoder	Finds unknown anomalies	May flag benign outliers	Emerging fraud discovery
Graph Analytics / GNN	Detects rings and linked abuse	Data engineering intensive	Mule networks and collusive fraud

Table 2. Representative Performance Pattern (Synthesized from Literature and Practice)

Approach	Precision	Recall	Operational Interpretation
Rule-only system	Moderate	Low to Moderate	Strong for known fraud, weak for adaptive attacks
Supervised ML	High	Moderate to High	Best for known fraud with labeled history
Sequence model	Moderate to High	High in behavior cases	Improves ATO and event-chain detection
Unsupervised anomaly	Low to Moderate	Variable	Useful for discovery and analyst triage
Hybrid layered system	High	High	Best balance of coverage and operational control

Table 3. Recommended Action Strategy by Risk Tier

Risk Tier	Illustrative Score Range	Recommended Action	Business Objective
Low	0.00–0.30	Approve silently	Maximize customer convenience
Medium	0.31–0.60	Step-up authentication or beneficiary cooling	Reduce false positives while controlling risk
High	0.61–0.80	Soft hold or manual review	Protect funds with limited friction
Critical	0.81–1.00	Hard decline and alert investigation	Prevent immediate fraud loss

12. Implementation Challenges and Limitations

Despite strong promise, AI-based fraud detection faces multiple implementation barriers. Data quality is one of the most common issues. Inconsistent labels, delayed dispute outcomes, missing telemetry, duplicate entities, and fragmented customer identity views can significantly reduce model effectiveness. Without reliable feature lineage and entity resolution, even advanced algorithms underperform.

Concept drift is another major challenge. Fraudsters adapt rapidly once detection controls become predictable. As payment platforms evolve, user behavior also changes due to seasonality, new product launches, promotional campaigns, and policy shifts. Models that perform well today may degrade silently over time. Therefore, drift monitoring, champion-challenger strategies, and scheduled plus event-driven retraining are necessary.

False positives remain a business-critical limitation. Excessive friction can reduce conversion, increase customer complaints, and damage trust in the payment platform. This is especially sensitive in instant payment systems where users expect real-time completion. Hence, institutions must balance security with usability through tiered responses such as step-up authentication, beneficiary cooling periods, and selective review instead of blanket blocking.

Privacy, fairness, and explainability also require attention. Payment data is highly sensitive, and institutions must ensure lawful data usage, access control, minimization, and retention governance. Models should be reviewed for unintended bias, especially where proxy variables may affect outcomes. Explainability tools and audit logs are essential for internal governance and external regulatory review.

Finally, many academic studies are limited by public datasets that do not fully represent real-world payment fraud complexity. As a result, there is a gap between laboratory performance and production effectiveness. This paper addresses that limitation by emphasizing system design, operational realism, and layered defense rather than overclaiming benchmark superiority.

13. Future Scope

Future research in digital payment fraud detection is likely to focus on graph-native and multimodal AI systems. As fraud becomes more coordinated, link analysis across devices, accounts, merchants, telecom signals, and external intelligence will become increasingly important. Graph neural networks, temporal graphs, and relation-aware embeddings can significantly improve detection of mule networks and collusive structures.

Federated learning and privacy-preserving machine learning may also become important in cross-institution settings where multiple banks or payment providers wish to benefit from shared intelligence without directly exchanging raw customer data. Techniques such as secure aggregation, differential privacy, and encrypted computation may help balance collaboration with compliance.

Another promising direction is adaptive decision intelligence that combines predictive models with reinforcement-learning-inspired policy optimization for intervention selection. Instead of only deciding whether a transaction is risky, future systems may optimize which action—approve, challenge, delay, or investigate—produces the best long-term outcome under customer and operational constraints.

Large language models (LLMs) may also contribute indirectly through investigator copilots, alert summarization, fraud pattern explanation, case triage assistance, and natural-language policy support. However, LLMs should be used carefully and not as standalone transaction decision makers in high-risk financial workflows without strict controls and validation.

14. Conclusion

This research paper examined AI-based fraud detection in digital payment systems through a structured academic framework aligned with contemporary financial technology risks. The study established that digital payment fraud is dynamic, adversarial, behavior-driven, and increasingly networked. Traditional rule-based systems remain useful for deterministic controls, but they are insufficient as a standalone defense in modern high-volume, real-time payment environments.

The paper demonstrated that the most effective strategy is a layered AI architecture combining supervised risk scoring, unsupervised anomaly detection, sequence modeling, graph intelligence, explainability, and feedback-driven model governance. It further highlighted the importance of feature engineering, real-time constraints, business-aware thresholding, and operational metrics beyond simple accuracy. The analysis showed that fraud prevention must be evaluated not only by detection quality but also by customer experience, latency, regulatory defensibility, and organizational scalability.

In conclusion, AI is not merely an enhancement for digital payment fraud control; it is becoming a foundational capability for secure and resilient payment infrastructure. Institutions that adopt responsible, explainable, continuously monitored AI frameworks will be better positioned to reduce fraud loss, maintain user trust, and support the continued expansion of digital financial services.

15. Acknowledgement

The authors would like to thank the faculty members, academic mentors, and institutional reviewers who support research in artificial intelligence, cybersecurity, and financial technology. The conceptual framework presented in this paper is intended for academic study and educational use, and it may be extended in future work using institution-specific datasets and experimental validation.

References

[1]. Bhattacharyya, S., Jha, S., Tharakunnel, K., & Westland, J. C. (2011). Data mining for credit card fraud: A comparative study. Decision Support Systems, 50(3), 602–613.

[2]. Dal Pozzolo, A., Caelen, O., Johnson, R. A., & Bontempi, G. (2015). Calibrating probability with undersampling for unbalanced classification. 2015 IEEE Symposium Series on Computational Intelligence, 159–166.

[3]. Dal Pozzolo, A., Boracchi, G., Caelen, O., Alippi, C., & Bontempi, G. (2018). Credit card fraud detection: A realistic modeling and a novel learning strategy. IEEE Transactions on Neural Networks and Learning Systems, 29(8), 3784–3797.

[4]. Jurgovsky, J., Granitzer, M., Ziegler, K., Calabretto, S., Portier, P. E., He-Guelton, L., & Caelen, O. (2018). Sequence classification for credit-card fraud detection. Expert Systems with Applications, 100, 234–245.

[5]. Fiore, U., De Santis, A., Perla, F., Zanetti, P., & Palmieri, F. (2019). Using generative adversarial networks for improving classification effectiveness in credit card fraud detection. Information Sciences, 479, 448–455.

[6]. Carcillo, F., Le Borgne, Y.-A., Caelen, O., Bontempi, G., & others. (2021). Combining unsupervised and supervised learning in credit card fraud detection. Information Sciences, 557, 317–331.

[7]. Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30.

[8]. 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 of literature. Decision Support Systems, 50(3), 559–569.

[9]. Phua, C., Lee, V., Smith, K., & Gayler, R. (2010). A comprehensive survey of data mining-based fraud detection research. arXiv preprint arXiv:1009.6119.

[10]. West, J., & Bhattacharya, M. (2016). Intelligent financial fraud detection: A comprehensive review. Computers & Security, 57, 47–66.

[11]. Abdallah, A., Maarof, M. A., & Zainal, A. (2016). Fraud detection system: A survey. Journal of Network and Computer Applications, 68, 90–113.

[12]. Le Khac, N.-A., Markos, S., Kechadi, T., & Le-Khac, N. (2020). Application of machine learning in fraud detection: A review. International Journal of Data Science and Analytics, 10(4), 1–15.

[13]. Roy, A., Sun, J., Mahoney, R., Alonzi, L., Adams, S., & Beling, P. (2018). Deep learning detecting fraud in credit card transactions. Systems and Information Engineering Design Symposium, 129–134.

[14]. Ryman-Tubb, N. F., 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.

[15]. Van Vlasselaer, V., Eliassi-Rad, T., Akoglu, L., Snoeck, M., & Baesens, B. (2017). GOTCHA! Network-based fraud detection for social security fraud. Management Science, 63(9), 3090–3110.

[16]. Randhawa, K., Loo, C. K., Seera, M., Lim, C. P., & Nandi, A. K. (2018). Credit card fraud detection using AdaBoost and majority voting. IEEE Access, 6, 14277–14284.

[17]. Bahnsen, A. C., Aouada, D., & Ottersten, B. (2016). Example-dependent cost-sensitive decision trees. Expert Systems with Applications, 42(19), 6609–6619.

[18]. Whitrow, C., Hand, D. J., Juszczak, P., Weston, D., & Adams, N. M. (2009). Transaction aggregation as a strategy for credit card fraud detection. Data Mining and Knowledge Discovery, 18, 30–55.

Appendix Note: Institutions may extend this paper by adding empirical dataset descriptions, confusion matrices, ROC/PR curves, ablation studies, and a deployment architecture diagram to align with dissertation or journal-specific page requirements.

Cite this Article:

Gholve, S. S., & Shinde, B. K. (2025). AI-based fraud detection in digital payment systems. International Journal of Emerging Research in Computer Science, 2(1), 1–9.

Journal URL: *https://ijercs.com/*

DOI: https://doi.org/10.59828/ijercs.v2i1.7