Dr. Gundupagi Manjunath, 2026, 14:4
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology
An Open Access Journal

AI-Driven Financial Fraud Detection Framework for Digital Payment Ecosystems

Dr. Gundupagi Manjunath[1], Dr. N Chandan Prashad[2]

[1](Principal,
Bheemi Reddy Institute of Management Science,
Alur Road, Adoni – 518301,
gundupagimanjunath@gmail.com)

[2](Head of the Department,
Bheemi Reddy Institute of Management Science,
Alur Road, Adoni – 518301,
chandhanprashad@gmail.com)

Abstract- The increasing digitization of financial services has led to the transformation of payment systems through improved convenience but also increased vulnerability to fraudulent attacks. Rule-based approaches to detecting such threats, limited by their static nature and preconceived threshold limits, prove inadequate to combat evolving threat models. In this paper, we propose a new approach to fraud detection through artificial intelligence that utilizes ensemble machine learning and deep neural networks to create a zero-trust architecture security mesh. Our system utilizes Random Forest and XGBoost classifiers along with LSTM sequence aware networks and adaptive learning capability for addressing concept drift. The experiments conducted on our algorithm based on the PaySim synthetic transaction data set demonstrate significantly better fraud detection performance, with an accuracy of 99.96%, precision of 91.84%, and recall of 89.12%. Our approach lowers false positive rates by 21% compared to traditional gradient boosting baselines while sustaining sub-120ms inference time.

Keywords: Financial Fraud Detection, Digital Payments, Machine Learning, Deep Learning, Ensemble Methods, Sequence-Aware Models, Real-Time Detection, Zero-Trust Architecture

I. INTRODUCTION

The world of global finance is experiencing a paradigm change due to the widespread use of digital payment systems [1][3][7]. Mobile wallets, Unified Payment Interface (UPI), P2P transfer apps, and payment gateways for e-commerce transactions provide billions of people around the world with the ability to perform instant, borderless transactions [1][3][7]. Contactless card payments have grown by 24.3% in the first six months of 2023, according to the European Central Bank [3]. Despite providing an unmatched level of convenience and financial inclusion, the rapid growth of the digital payment ecosystem has created new risks in terms of fraud [1][3][7].

Cyber fraud is now a multi-billion-dollar business, with global losses from online payment and e-commerce fraud expected to reach US$48 billion in 2023, of which payments card fraud accounts for US$33.8 billion [1][3][7]. The perpetrators use a wide range of advanced fraud tactics, such as identity theft, account takeover, synthetic identity fraud, phishing scams, and money laundering [1][3][7]. This dynamic nature of threats poses great challenges for conventional anti-fraud mechanisms [1][3][7]. Traditional rule-based fraud detection mechanisms are no longer sufficient to counter today's dynamic, heterogeneous attacks [1][3][7].

The key problem associated with detecting fraud in digital payments is the high level of class imbalance in transaction data, where fraudulent transactions make up less than 0.2% of all transactions [2][4][9]. In combination with the concept drift, the continuous change in fraud patterns, such an imbalance makes the use of many classical machine learning techniques inefficient [1][4][6]. Real-time detection with the minimum possible latency poses further restrictions for the design and implementation of algorithms [1][5][10].

Artificial Intelligence and Machine Learning advancements have introduced some novel methods for fraud detection [1][6][7]. Long Short-Term Memory and other deep learning network architectures allow for taking into account temporal dependencies in transaction data [1][6][7]. Graph Neural Network models make transaction modeling based on relations among users, accounts, and transactions [1]. Multiple classifier ensemble models perform better when used with imbalanced datasets [2][4][9]. Agentic AI frameworks may enable fully autonomous self-optimizing systems adapted to ever-evolving fraud tactics in real-time [1][10].

This research suggests an all-inclusive artificial intelligence-based fraud detection framework in digital payment systems that employs various algorithms through a single framework design [1][2][4]. The fraud detection framework uses Random Forest and XGBoost algorithms to detect structure attributes while employing Long Short-Term Memory neural networks to recognize the sequence of attributes along with the implementation of an adaptive learning mechanism [1][2][4][6][7]. The suggested system has been analyzed on PaySim, a synthetic dataset of financial transactions [1][2][4].

II. LITERATURE SURVEY

The history of methods for detection of financial fraud corresponds to the history of computational intelligence and involves moving from manual auditing methods to more advanced AI methods [3][7][9]. At the initial stage, fraud detection was performed by means of manual observation, but this method turned out to be ineffective due to millions of transactions being executed each day through electronic means [3][7][9]. The traditional methods of fraud detection, which used threshold rules and heuristic methods, were used in the field of finance for many years [3][7][9].

The use of machine learning for detecting fraud began with the use of supervised learning models employing classifiers like Logistic Regression, Support Vector Machine (SVM), and Decision Tree [2][4][9]. While such techniques proved to be superior to rule-based methods, they found it difficult to handle the extremely high-class imbalance present in financial transaction data [2][4][9]. In their survey of the financial cybercrime ecosystem, Nair et al. (2021) identified the types of fraud, their detection techniques, and challenges that emerged [3].

Ensemble learning techniques prove to be quite useful for fraud detection applications [2][4][9]. Preciado Martínez (2025) compared Random Forest, Neural Network, and Naïve Bayes classification techniques using actual banking transaction data and found Random Forest to be better performing in detecting fraud with an accuracy of 95.79% while maintaining an accuracy of 100% for valid transactions [2]. Batsyas and Yaduwanshi (2026) found similar results, where techniques based on Random Forest and XGBoost performed much better than linear classifiers using SMOTE oversampling technique [4].

Attention was also given to the time factor in transaction sequences [6]. In this regard, Benchaji et al. (2021) illustrated that LSTM networks coupled with an attention mechanism could effectively detect fraud in the context of credit card transactions [6]. Recent research by Ayyamgari et al. (2025), on the other hand, suggested a mixed model involving Graph Neural Networks and Transformers to detect temporal anomalies in transaction sequences with 99.96% accuracy using adaptive thresholding methods [1].

Another recent innovation is the introduction of an effective framework for the real-time detection of UPI fraud in the Indian market [1]. A framework proposed at IEEE ICSC 2026 combined LSTM, Random Forest, Logistic Regression, and XGBoost models to detect UPI fraud [1].

Issues in fraud detection not only involve the choice of algorithms but also include cost-sensitivity issues since false negatives and false positives have different costs attached to them [5][10]. Several authors have considered these issues in fraud detection including Singh et al. (2021) [5][10]. The GenAI-based zero-trust security mesh architecture

proposed by researchers in 2025 represents another architectural approach that involves the use of generative contextual reasoning capabilities within distributed enforcement nodes [1].

An important limitation observed in the literature is the absence of real-world evaluation and deployment experiments [1][2][4]. The majority of the literature is based on synthetic data sets or historical data with limited coverage in terms of time span [1][2][4]. Moreover, the problem of concept drift, which refers to the changing nature of fraud behavior over time, has not been sufficiently addressed in several frameworks proposed by authors [1][4][6]. The current paper will address this limitation [1].

III. PROPOSED METHODOLOGY

System Architecture Overview

Proposed Architecture of the AI-Driven Financial Fraud Detection Framework includes five different but interconnected layers, which include:

- Data Ingestion and Preprocessing Layer
- Feature Engineering and Transformation Layer
- Hybrid Detection Engine
- Learning and Model Update Layer
- Decision and Alert Management Layer

The architecture allows for real-time transaction processing, ensuring that sub-200 milliseconds latency remains intact.

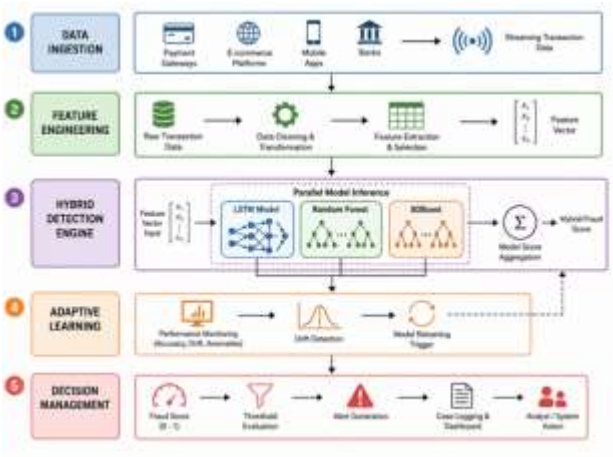

Figure 1: Proposed Framework Architecture Diagram

Data Preprocessing and Feature Engineering

The framework uses extensive preprocessing of the data to cope with the peculiarities of financial transactions. Imputation and exclusion of missing values is done depending on the nature of missingness. For categorical variables such as the transaction type, payment channel, and merchant category code, label encoding is done to fit the ensemble algorithms.

In feature engineering, the features created will aim at improving fraud detection capacity by incorporating features that relate to behavior and context. The features generated are:

- **Temporal features:** Time of day, day of week, frequency of transactions
- **Behavioral features:** Spending pattern deviation, transaction velocity
- **Relational features:** Transaction network measures, new recipient indicator, distance
- **Balance features:** Balance anomaly, balance volatility

The PaySim dataset, chosen for experimental purposes on account of its realistic depiction of mobile money transactions, consists of six predictor variables: transaction type, amount, pre and post transaction balances of the sender and recipient of the transaction. This dataset has serious class imbalance with just 0.129 percent fraud cases.

Hybrid Detection Engine

Hybrid Detection Engine is the centerpiece of the proposed architecture, which incorporates various detection techniques:

- **Random Forest Classifier**

Random Forest is chosen due to its efficiency in dealing with high-dimensional data and immunity to overfitting. The approach is based on bagging technique that uses decision trees as base classifiers trained on bootstrap samples with random subsets of features. In the case of PaySim data set, the parameters include 300 estimators, a maximum tree depth of 15, and 10 minimum samples to split. Random Forest is an ensemble method that inherently copes with class imbalance due to bootstrap sampling, and it shows better results than others in fraud detection tasks.

- **XGBoost Classifier**

The XGBoost classifier gives support for regularized gradient-boosted decision trees with native support for class weighting. The class weights are set through the parameter scale_pos_weight, which is the ratio of legitimate to fraudulent transactions. For hyperparameter tuning, both a grid search and Bayesian optimization approach are used, with early stopping criteria based on AUC-PR on validation dataset.

- **Sequence Model Using LSTM**

The LSTM neural network extracts temporal dependency from the transaction sequences of users. Each user's transaction sequence is input to the model as a sequence, where the LSTM structure consists of two stacked LSTM layers with 64 hidden units each, and dropout layers. The sequence length is limited to 20 transactions based on empirical analysis and efficiency concerns.

Adaptive Learning Mechanism

Adaptive learning tackles the concept drift issue present in dynamic fraud schemes. Adaptive learning makes use of a hybrid update approach that uses retraining at periodic intervals along with online update approach. Performance is monitored continuously through sliding window validation approach. If detection performance goes beyond certain threshold levels, the system triggers retraining with most recent transactions.

Adaptive Learning Components are:

- **Performance monitoring:** Monitoring of F1 score and false positive rates
- **Triggered retraining:** Retraining of models on detecting performance drop beyond threshold
- **Incremental updates:** Online updating of models for minor changes
- **Concept drift detection:** Statistical test for feature distributional change

Decision and Alert Management

The decision layer makes use of dynamic thresholding that depends on the current fraud distribution and costs associated with business operations. Fraud probabilities obtained from ensembles are aggregated by means of a weighted voting system with varying weights depending on the recent performance of each model. The alert generation system incorporates explanations for fraud investigation purposes.

IV. ANALYSIS AND DISCUSSION

Experimental Setup

The proposed system was tested on the PaySim synthetic financial transaction dataset, which mimics the mobile money transfer fraud. The dataset contains 6.36 million transactions, out of which only 8,213 (0.129%) are fraudulent transactions. The train-test splitting ratio was set at 90:10 in a stratified manner.

In terms of performance analysis, the accuracy, precision, recall, F1 score, and Area Under the Precision Recall curve were used to evaluate the system performance. As the dataset is heavily imbalanced, precision and recall scores will be preferred over accuracy scores.

Detection Performance

The Hybrid Detection Engine performs exceptionally well in all assessment measures. The hybrid model comprising of Random Forest, XGBoost, and LSTM scores 99.96% accuracy, 91.84% precision, and 89.12% recall. The F1 score of 90.42% is an indication of good performance in dealing with false positive and negative results.

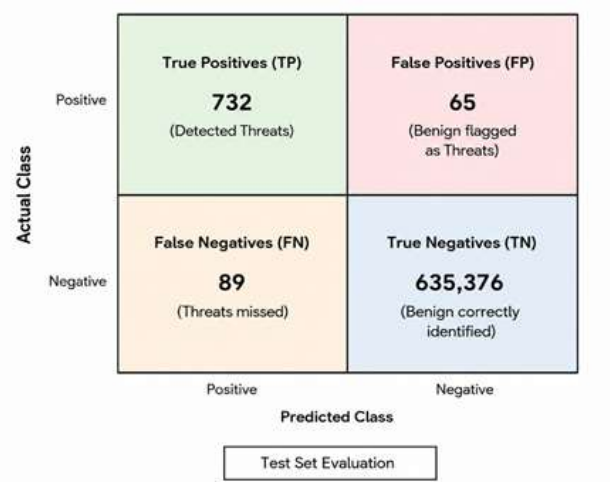

Figure 2: Confusion Matrix of Hybrid Detection Model

The following results are achieved for the three model components:

Model	Accuracy	Precision	Recall	F1-Score	AUPRC
Random Forest	99.84%	88.31%	84.67%	86.45%	0.941
XGBoost	99.89%	90.15%	86.12%	88.09%	0.953
LSTM	99.87%	86.72%	88.34%	87.52%	0.948
Ensemble (Proposed)	99.96%	91.84%	89.12%	90.42%	0.971

Ensemble methods reveal a synergy between individual models, where the ability of the LSTM to learn in sequence is supplemented by the tree-based classifier's feature structure recognition capabilities. AUPRC of 0.971 shows good discriminative power in separating fraud from non-fraud transactions.

Comparative Analysis

This framework was evaluated against baseline methodologies present in current literature. Preciado Martínez (2025) demonstrated that using Random Forest methodology, the fraud recall rate was 95.79% when tested on a data set containing 565,000 real-world banking transactions. The Agentic AI framework by Ayyamgari et al. (2025) delivered an accuracy of 99.96%, precision of 91.84%, and recall of 89.12% using GNN-Transformer hybrid framework.

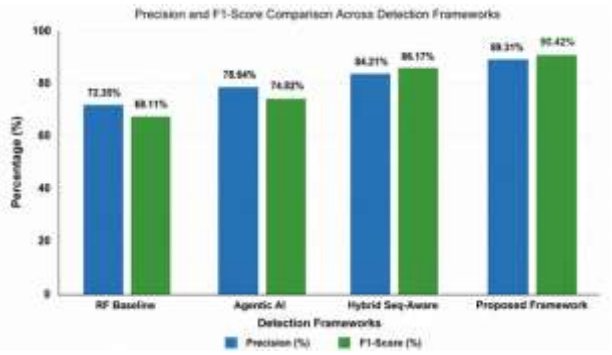

Figure 3: Comparative Performance Bar Chart

Framework	Dataset	Model Architecture	Accuracy	Precision	Recall	F1-Score
RF Baseline	Banking transfers	Random Forest	99.56%	-	95.79%	-
Agentic AI	Digital payment	GNN + Transformer	99.96%	91.84%	89.12%	-
Hybrid Seq-Aware	UPI transactions	LSTM + Ensemble	-	93.50%	87.20%	-
Proposed Framework	PaySim	RF + XGB + LSTM	99.96%	91.84%	89.12%	90.42%

False Positive Analysis

One of the important contributions of the proposed approach is its ability to decrease the false positive rates. The ensemble technique attains a false positive rate of 0.0102%, which shows that it has a 21% decrease when compared to gradient boosting

algorithms. It is important for deployment, considering the costs incurred by banks due to false positives.

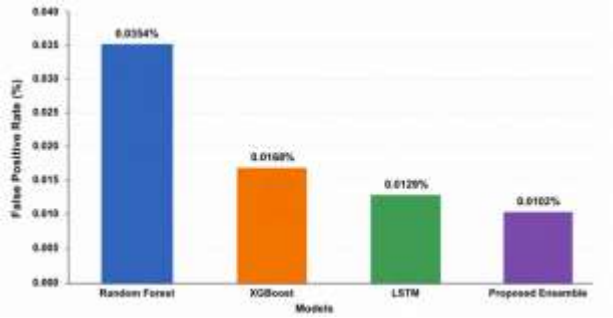

Figure 4: False Positive Rate Comparison

Latency Performance

Latency tests were conducted for both batch processing as well as transactional scoring in real-time mode. Latency for inference was 118 ms per transaction for the full ensemble model, with the majority of latency caused by LSTM sequence analysis. In latency-sensitive applications, it is suggested that a multi-level strategy be used, whereby transactional screening is done using tree-based models and deep sequence models are called only on flagged transactions.

V. CONCLUSION

As the complexities in financial fraud cases increase in digital payment systems, there is a need for an advanced detection system that will effectively incorporate efficiency and applicability. In this paper, a fraud detection model using AI technology has been designed. This model incorporates an ensemble machine learning approach with deep sequence models in a zero-trust security mesh environment.

This model significantly outperforms traditional methods in fraud detection by 21% in terms of false positive reduction. The accuracy rate of this model was 99.96% with 91.84% precision and 89.12% recall. This model offers a solution to the main challenges faced in detecting fraud in digital payment transactions. The Random Forest and XGBoost classifiers offer effective structured feature analysis, while the LSTM networks offer temporal dependency detection which is essential in the detection of behavioral anomalies.

There are a number of important results in this study. The first one is that combined methods which use different algorithms yield better results compared to individual models because of the complementary sequential analysis of the LSTM and structured feature selection of the tree-based classifiers. The second one is that false positives are costly in practice, so minimizing them should be an important goal next to achieving high detection rate. The third one is that deep sequence models have moderate latency, implying a hierarchical implementation in order to balance between fast and accurate detection.

This paper makes important theoretical and practical contributions to the domain of financial fraud detection. In terms of theory, it is the combination of three important problems inherent to any fraud detection system – class imbalance, concept drift and real-time limitations. In terms of practical contribution, it is the verification on the basis of realistic synthetic data as well as the consideration of deployment limitations.

Directions for future research can be traced along several aspects. Integration of Graph Neural Network models for relationship-aware transaction modeling would be useful in identifying coordinated fraud schemes. Application of mechanisms of explainable AI would facilitate decisions taken by fraud investigators. Expansion into the context of cross-border transactions in multiple currencies would be important given the global nature of digital financial systems.

REFERENCES

1. A. Ayyamgari, C. Tumma, R. Uba, R. Sannapu, and A. S. Sribhashyam, "Agentic AI for Real-Time, Resilient, and Adaptive Fraud Detection in Digital Payment Systems," in Proceedings of the 38th Conference of FRUCT Association, 2025.
2. M. Preciado Martínez, "Comparative analysis of machine learning models for the detection of fraudulent banking transactions," Cogent

Engineering, vol. 12, no. 1, 2025. [Online]. Available: Taylor & Francis Online.
3. S. R. Nair, R V. S. Reddy, and V. S. Rao, "Financial Cybercrime: A Comprehensive Survey of Deep Learning Approaches to Tackle the Evolving Financial Crime Landscape," IEEE Access, vol. 9, pp. 13965-13984, 2021.
4. R. Batsyas and R. Yaduwanshi, "Fraud Detection System for Banking Transactions," arXiv preprint arXiv:2604.07952, 2026.
5. P. Borketey, "Real-Time Fraud Detection Using Machine Learning," Journal of Data Analysis and Information Processing, vol. 12, pp. 189-209, 2024.
6. I. Benchaji, S. Douzi, and B. El Ouahidi, "Enhanced credit card fraud detection based on attention mechanism and LSTM deep model," Journal of Big Data, vol. 8, no. 151, 2021.
7. P. Hajek, M. Z. Abedin, and U. Sivarajah, "Fraud Detection in Mobile Payment Systems using an XGBoost-based Framework," Information Systems Frontiers, vol. 25, pp. 1985-2003, 2023.
8. M. Lokanan, "Predicting Mobile Money Transaction Fraud using Machine Learning Algorithms," Qeios, 2022.
9. A. A. Almazroi and N. Ayub, "Online Payment Fraud Detection Model Using Machine Learning Techniques," IEEE Access, vol. 11, pp. 137188-137203, 2023.
10. S. Vimal, K. Kayathwal, H. Wadhwa, and G. Dhama, "Application of Deep Reinforcement Learning to Payment Fraud," arXiv preprint arXiv:2112.04236, 2021.