

Artificial Intelligence-Based Fraud Detection in Financial Networks: A Comprehensive Review of AI Techniques

Nikita Vikas Chavan¹, Prof. S. S. Medhe², Dr. H. B. Jadhav³

Student, Department of Computer Engineering¹

Guide, Department of Computer Engineering²

Co-Guide, Department of Computer Engineering³

Adsul's Technical Campus, Chas Ahilyanagar, India

Abstract: *The rapid growth of digital banking, card payments, mobile wallets and online financial services has created large and continuously changing transaction networks. Fraudulent behaviour can appear as abnormal amounts, unusual transaction sequences, account takeover, identity misuse, suspicious device changes, coordinated merchant activity and hidden relationships among entities. Conventional rule-based systems remain useful for clearly known patterns, but they require frequent rule maintenance and can generate a high number of false alerts. This paper provides a comprehensive review of artificial intelligence-based fraud detection techniques used in financial environments. Machine learning, deep learning, anomaly detection, ensemble learning and graph-oriented methods are considered together with data preprocessing, feature engineering, class-imbalance handling, explainability, privacy and continuous monitoring. The paper also describes an existing rule-based workflow and a proposed AI-assisted workflow in which transaction data are cleaned, behavioural and relational features are extracted, suitable AI models generate fraud-risk scores, explanations support investigation, and feedback is used for monitoring and controlled model improvement. The review is literature-based and conceptual; therefore, it does not claim a new measured accuracy result. Dataset-specific experimental validation is required before operational deployment.*

Keywords: Artificial intelligence, fraud detection, machine learning, deep learning, financial networks, anomaly detection, explainable AI, class imbalance, concept drift, federated learning.

I. INTRODUCTION

Financial institutions process large numbers of transactions every day through interconnected banking and payment channels. The same connectivity that improves convenience also creates opportunities for fraudulent activity to move rapidly between accounts, devices, merchants and services. A fraud detection system therefore has to examine both individual transaction attributes and the wider context in which a transaction occurs.

Traditional detection mechanisms commonly use expert-defined rules, thresholds, blacklists and manual review. Examples include unusually large amounts, repeated transactions within a short period, impossible location changes and transactions associated with previously identified indicators. These methods can be transparent and operationally simple, but their effectiveness depends on how quickly new rules are created and how accurately the rules represent current fraud behaviour.

Recent research has investigated supervised machine learning, unsupervised anomaly detection, deep neural networks, hybrid architectures and graph-based approaches. Sarna et al. [1] review AI, ML, DL and hybrid methods for financial



fraud. Almazroi and Ayub [2] describe an online-payment fraud model using preprocessing, imbalance handling and deep sequence modelling. Hilal et al. [4] review anomaly-detection techniques and related challenges. This review considers AI techniques together with the data and operational conditions that affect their use. The discussion covers model families, data preparation and class imbalance, evaluation, explainability, privacy and continuous monitoring. Results from different datasets are not treated as directly comparable because fraud prevalence, features, labels, preprocessing and evaluation protocols can substantially affect measured performance.

II. PROPOSED SYSTEM

The proposed system extends the conventional workflow by inserting data-quality processing, feature extraction, AI-based prediction, risk scoring, explainable alerts and controlled feedback. The detection layer can use supervised classification, anomaly detection, deep learning, graph analysis or an ensemble according to the available data.

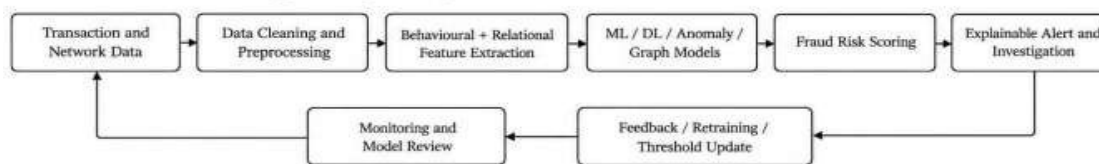


Fig. 1. Proposed System – AI-Based Fraud Detection Workflow

The workflow starts with transaction and network data. Preprocessing validates records and creates consistent model inputs. Feature extraction combines transaction attributes with behavioural indicators such as frequency, velocity and deviation from historical patterns. Relational features can represent connections among accounts, devices, merchants and other entities.

The detection layer can use supervised machine learning when representative labels are available, deep learning for complex patterns, anomaly detection when labels are limited, and graph methods when relationships provide useful evidence. The final risk score should be interpreted using a threshold selected from validation data.

Explainable alerts provide analysts with factors or relationships that contributed to a high-risk decision. Investigation outcomes can be stored as controlled feedback. Monitoring should track input distributions, alert rates, model behaviour and confirmed outcomes so that recalibration or retraining can be considered when drift is detected.

III. EXISTING SYSTEM

The conventional system captures a transaction, applies predefined rules or thresholds, flags suspicious records and moves them to manual verification. It is understandable and useful for known patterns, but its adaptability depends on rule updates, and complex fraud may require many interacting rules.

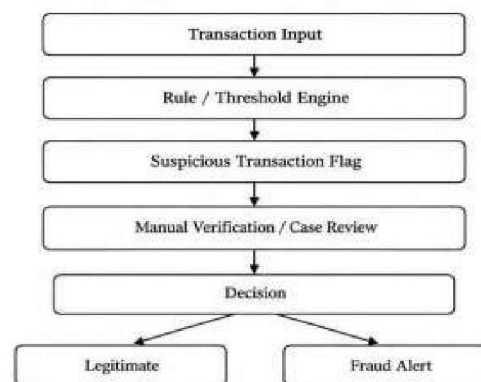


Fig. 2. Existing System – Conventional Rule-Based Fraud Detection



The main limitation is that detection logic is primarily defined in advance. A new fraud pattern may remain undetected until an appropriate rule or indicator is identified and deployed. Rule-based controls can therefore remain useful as one layer of a broader system

IV. METHODOLOGY

The methodology follows six connected stages: data collection, preprocessing, feature extraction, model training, risk scoring and continuous evaluation. The objective is to ensure that evaluation conditions resemble actual deployment rather than relying only on a random split that may mix information from different time periods.

a) Dataset Preparation

Fraud records are typically much fewer than legitimate records. The dataset should therefore be inspected for class distribution, missing values, duplicate transactions and temporal patterns. Resampling, synthetic generation or class-weighting can be applied within training. The reviewed online-payment study uses SMOTE as one approach [2]. Transformations that learn from data should be fitted only on the training portion to reduce data leakage.

b) Training and Validation

The data can be separated into training, validation and test sets. When transaction behaviour changes with time, a chronological split is preferable because it reduces the risk of using future information to predict the past. Hyperparameters should be selected using validation data, while the test set remains isolated until final evaluation.

c) Risk Threshold

The model output is converted into an operational decision using a threshold. A lower threshold may identify more potential fraud but can increase false alerts, while a higher threshold can reduce alerts but may miss some fraud. Threshold selection should use validation results and the operational cost of different errors.

V. PERFORMANCE EVALUATION

Because fraud is commonly an imbalanced classification problem, evaluation should not depend on accuracy alone. Accuracy, precision, recall, F1-score, ROC-AUC and PR-AUC describe different aspects of detection quality.

a) Accuracy

Accuracy is the proportion of all correct predictions. It can be misleading for highly imbalanced fraud data because a model may obtain high accuracy by predicting most records as legitimate.

b) Precision

Precision measures the proportion of predicted fraud cases that are actually fraudulent. It is relevant when analyst investigation capacity is limited and excessive false alerts are costly.

c) Recall

Recall measures the proportion of actual fraud cases that are detected. It is important when missed fraud creates substantial financial or operational impact.

d) F1-Score

F1-score combines precision and recall into a single measure and is useful when both false positives and false negatives matter.

e) ROC-AUC and PR-AUC

ROC-AUC summarises discrimination across thresholds. Precision-recall analysis can be especially informative when the positive fraud class is rare because it focuses on precision and recall for the minority class.

A complete evaluation can also report false-alert rate, processing latency and model stability across time. This paper does not report a new experimental score because no common dataset and controlled experiment were provided.



VI. COMPARATIVE ANALYSIS

The reviewed literature covers several major families of fraud-detection approaches. Their characteristics differ in evidence type, data requirements and operational complexity. The comparison below is qualitative and summarises methodological characteristics rather than ranking the approaches.

Approach	Main Strength	Main Limitation
Rule-based	Transparent and simple	Fixed logic; maintenance required
Machine learning	Learns from historical features	Needs representative labels and features
Deep learning	Models complex non-linear patterns	Higher compute and explain ability requirements
Anomaly detection	Can discover unusual behaviour	Rare legitimate events may trigger alerts
Graph/hybrid	Captures relational and combined evidence	More complex data and deployment

Table 1. Qualitative comparison of major fraud-detection approaches discussed in the reviewed literature. Rule-based methods remain useful for known conditions. Machine-learning methods learn from historical features, deep models can represent complex patterns, anomaly detection can highlight unusual behaviour, and graph/hybrid methods can add relational evidence. Each approach has data, interpretability or deployment considerations.

VII. CLASS IMBALANCE AND CONCEPT DRIFT

Fraud detection has two linked data problems: fraud is usually a minority class, and behaviour can change over time. Class imbalance can bias a classifier toward legitimate transactions, while concept drift can make historical training data less representative of current behaviour.

a) Class Imbalance

Possible responses include class weighting, controlled undersampling, oversampling, synthetic generation and anomaly-based modelling. The selected method should be evaluated using precision, recall, F1-score and precision-recall analysis rather than accuracy alone.

b) Concept Drift

Drift can occur when fraudsters change tactics, customers change payment behaviour, or new channels alter transaction distributions. Monitoring can compare production inputs and outcomes with training distributions; significant changes can trigger investigation, recalibration or retraining.

VIII. EXPLAINABILITY AND PRIVACY

Explainability is important because fraud alerts are often reviewed by analysts. A useful explanation can identify features or relationships that contributed to a high-risk score and supports auditing. Privacy is equally important because transaction records contain sensitive financial information. Data minimisation, role-based access, encryption, secure logging and controlled retention reduce unnecessary exposure. Federated learning can allow local model training while reducing the need to pool raw transaction records [3].

IX. SECURITY AND REAL-TIME PROCESSING

A production fraud detector must operate under latency and reliability constraints. Incoming transactions should pass through validation, feature generation and model inference within an acceptable response time. Fallback behaviour should be defined for model failures, missing features and uncertain predictions.



Attackers may adapt to detection logic or manipulate input attributes. Monitoring should include unusual feature distributions, sudden alert-volume changes, repeated attempts from connected entities and unexpected model-output patterns. Model versions, thresholds and rule changes should be logged for auditability, and access to training data, model artefacts and investigation records should be controlled.

X. DISCUSSION

The reviewed studies demonstrate that AI-based fraud detection is not limited to choosing one classifier. The overall pipeline determines practical usefulness. Data preparation affects learned patterns; class imbalance affects training; threshold selection affects alert workload; explainability affects investigation; and monitoring affects performance after deployment.

A hybrid architecture can combine complementary signals. A supervised model may provide a primary risk score, an anomaly detector may highlight unusual behaviour, and a graph model may add relational evidence. The final decision can be supported by explanation and human review. Such combinations should be validated experimentally rather than assumed to provide an automatic improvement.

XI. EXPECTED BENEFITS OF THE PROPOSED SYSTEM

The proposed system separates data preparation, detection, explanation and monitoring. It can accommodate multiple model families, produce risk-based rather than purely rule-based decisions, reduce dependence on manual rule creation and provide a structured feedback path. These are design benefits and should not be interpreted as measured performance improvements.

XII. LIMITATIONS

This paper presents a literature-based review and a conceptually proposed system. It does not claim a new measured accuracy result for a particular financial dataset. Reported performance from different papers cannot be compared directly because datasets, fraud prevalence, features, preprocessing and evaluation protocols differ. Practical implementation requires controlled experiments on representative transaction data.

The review does not establish that one AI technique is universally suitable for all financial networks. Model selection depends on label quality, temporal structure, relational information, privacy constraints, computational resources and operational requirements.

XIII. FUTURE SCOPE

Future work can implement and compare several model families on the same dataset and under the same temporal evaluation protocol. Research can also investigate graph neural networks, transformer-based transaction models, federated learning, privacy-preserving aggregation and explainable hybrid systems. Real-time streaming pipelines can be evaluated using both detection quality and inference latency.

Another direction is human-in-the-loop learning. Analyst decisions can be stored as labelled feedback, reviewed for quality and incorporated into controlled retraining cycles. Drift detection can identify when recalibration is needed. Future studies should report precision, recall, F1-score, PR-AUC, false-alert rate, processing latency and model stability across time.

Future experiments should document dataset construction, class distribution, preprocessing, feature definitions, train/validation/test separation, threshold selection and reproducibility settings. Such reporting would make comparisons between AI techniques more meaningful.

XIV. CONCLUSION

This paper reviewed artificial intelligence-based fraud detection techniques for financial networks, with emphasis on machine learning, deep learning, anomaly detection, ensemble approaches and graph-oriented analysis. Traditional



rule-based detection remains useful for known conditions, while changing fraud behaviour motivates data-driven methods that incorporate behavioural and relational evidence.

The proposed system begins with transaction and network data, performs preprocessing and feature extraction, applies suitable AI techniques, generates a fraud-risk score and produces an explainable alert for investigation. Monitoring and feedback provide a controlled path for detecting drift and improving future model versions. The workflow connects detection techniques with practical requirements such as privacy, interpretability, scalability and real-time operation.

The main contribution of this review is the organisation of AI techniques and operational requirements into a coherent fraud-detection workflow without claiming a new experimental result. Dataset-specific implementation and evaluation are necessary before conclusions about operational performance can be made.

XV. ACKNOWLEDGMENT

The authors acknowledge the academic guidance and research support provided by the department and institution. The authors also acknowledge the publicly available research literature used to structure this review and the proposed fraud-detection workflow.

REFERENCES

- [1] N. J. Sarna, F. A. Rithen, U. S. Jui, S. Belal, Al Amin, T. K. Oishee, and A. K. M. Muzahidul Islam, "AI Driven Fraud Detection Models in Financial Networks: A Comprehensive Systematic Review," *IEEE Access*, vol. 13, pp. 141204–141233, 2025, doi: 10.1109/ACCESS.2025.3596060.
- [2] A. A. Almazroi and N. Ayub, "Online Payment Fraud Detection Model Using Machine Learning Techniques," *IEEE Access*, vol. 11, pp. 137188–137203, 2023, doi: 10.1109/ACCESS.2023.3339226.
- [3] T. Awosika, R. M. Shukla, and B. Pranggono, "Transparency and Privacy: The Role of Explainable AI and Federated Learning in Financial Fraud Detection," *IEEE Access*, vol. 12, pp. 64551–64560, 2024, doi: 10.1109/ACCESS.2024.3394528.
- [4] W. Hilal, S. A. Gadsden, and J. Yawney, "Financial Fraud: A Review of Anomaly Detection Techniques and Recent Advances," *Expert Systems with Applications*, vol. 193, 116429, 2022, doi: 10.1016/j.eswa.2021.116429.
- [5] E. W. T. Ngai, Y. Hu, Y. H. Wong, Y. Chen, and X. Sun, "The Application of Data Mining Techniques in Financial Fraud Detection: A Classification Framework and an Academic Review of Literature," *Decision Support Systems*, vol. 50, no. 3, pp. 559–569, 2011.
- [6] D. Jurgovsky et al., "Sequence Classification for Credit-Card Fraud Detection," *Expert Systems with Applications*, vol. 100, pp. 234–245, 2018.
- [7] S. West and J. Bhattacharya, "Intelligent Financial Fraud Detection: A Comprehensive Review," *Computers & Security*, vol. 57, pp. 47–66, 2016.
- [8] IEEE, "IEEE Editorial Style Manual," IEEE Publication Services and Products Board, IEEE.
- [9] A. Al-Hashedi and Y. Magalingam, "Financial Fraud Detection Applying Data Mining Techniques: A Comprehensive Review from 2009 to 2019," *Journal of Applied Science and Engineering*, vol. 24, no. 6, pp. 787–804, 2021.
- [10] A. A. A. El-Sayed, M. M. H. Shalaby, and M. A. El-Bialy, "Credit Card Fraud Detection Using Machine Learning Techniques: A Review," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 6, 2022.

