Global Journal of Engineering and Technology Advances
eISSN: 2582-5003
Cross Ref DOI: 10.30574/gjeta
Journal homepage: https://gjeta.com/

(RESEARCH ARTICLE)

Check for updates

Knowledge extraction from financial network data for fraud and risk detection

Divine U Linus *

Bank of America – Global Risk Management, Financial Crimes.

Global Journal of Engineering and Technology Advances, 2026, 27(01), 074-087

Publication history: Received on 13 February 2026; revised on 20 March 2026; accepted on 23 March 2026

Article DOI: https://doi.org/10.30574/gjeta.2026.27.1.0066

Abstract

Financial fraud is a significant threat to financial institutions and the financial regulatory body because fraudulent transactions are difficult to detect and trace. The proposed research paper presents a viable artificial intelligence model for deriving knowledge from financial network data to support fraud and risk detection. This paper introduces a node-level financial risk scoring framework with precision-driven optimization tailored for regulatory fraud detection environments. The method combines graph-based structural characteristics and machine learning models to detect suspicious transactions and concealed network characteristics related to fraud. An Elliptic node-level risk indicator based on structural influence derived from network analysis of the Elliptic Bitcoin transaction data was used to build the predictive models. Results show that Node2Vec +XGBoost XGBoost and LightGBM have AUCs of 0.952, 0.945 and 0.938, respectively. Besides, LightGBM has high precision (0.975) and a very small false-positive rate (0.0013). These findings indicate that the framework can identify illegal transactions and reduce false alarms under the regulations. In addition, one uses graph insights to determine structurally powerful nodes that act as intermediaries in problematic transaction networks. Explainable Artificial Intelligence (XAI) can also increase transparency by allowing financial authorities and compliance groups to obtain risk ratings. Overall, the given framework demonstrates that the combination of graph analytics and explainable machine learning has the potential to facilitate the identification of financial fraud, risk monitoring, and regulatory decision-making in the contemporary financial landscape.

Keywords: Interpretable AI; Machine learning; Financial fraud detection; Graph-based network

1. Introduction

The rapid development of internet financial services and the introduction of online banking systems into the environment have dramatically transformed the structure of a modern financial system. Finance operations are currently conducted across a multi-layered network of clients, financial institutions, digital platforms, and payment systems. Although this digital transformation has led to financial efficiency and accessibility, it has also increased vulnerability to sophisticated financial fraud schemes and systemic risks. Credit card fraud, identity theft, and other transactions that tamper with transactions have become harder to trace as the volume of data recorded in financial transactions is very high and increasingly sophisticated. Consequently, financial institutions and regulators are turning to artificial intelligence (AI) and advanced analytics as increasingly attractive means of fraud detection and to make the process of financial risk management more efficient [1, 2].

Rule-based mechanisms and statistical anomaly detection models were the main elements of traditional fraud detection systems. Even though such methods may be used to detect established fraud patterns, they struggle to detect new and emerging ones. Current research shows that machine learning methods play a major role in detecting fraud because they are used to detect complicated behavioral patterns on financial data. The algorithms of machine learning will be able to identify the pattern of all transactions historically and identify the anomalies related to suspicious behavior, facilitating the effectiveness and precision of fraud detection systems [3]. Moreover, deep learning models are proving

* Corresponding author: Divine U Linus.

highly predictive for identifying anomalies in financial transaction streams because they can capture nonlinear associations in large-scale data [4].

However, because contributions to relational dependencies within financial entities are very strong, traditional machine learning models will most of the time treat them as independent observations. The financial transactions are, in fact, systems of interconnected networks in which accounts, customers, and institutions are in various financial relations. The interactions build intricate financial networks which possess useful structural information on fraudulent activity and the spread of systemic risk. Graph-based machine learning algorithms have thus attracted significant interest in the study of relational financial data. GNNs are especially applicable to financial network modeling, since this type of neural network will capture structural connections between entities and expose fraud patterns that are hidden in network structures [5, 6] .

State-of-the-art research has shown that graph neural networks (GNNs) can be used to detect financial fraud. Graph learning models can identify complex fraudulent networks by analyzing connections between financial entities. As an example, adaptive GNN models have been suggested to identify transaction fraud by adding network connectivity and behavioral features to predictive models [7]. Similarly, a graph convolutional network has been used to identify fraudsters in financial transactions by highlighting influential nodes in financial networks [8]. Additional frameworks that other researchers have suggested to facilitate more complex transaction relationships and boost the accuracy of fraud detection include encoder–decoder GNNs and metapath-guided graph learning models [9, 10].

Temporal and causal GNNs are another development that has been applied in the analysis of financial networks. Financial transaction networks are dynamic, and fraud acts tend to change over time. Models of temporal graph learning can learn time-evolving relationships between transactions and identify new patterns of fraud in changing financial systems [11].In similar fashion, causal temporal GNNs have also been created to detect causal relationships between financial transactions and enhance the detection of organized fraud patterns [12]. Other articles suggest using global confidence-based graph learning models to enhance fraud detection by incorporating network-level trust measurements into predictive systems [13].

Although these AI-based solutions have enhanced fraud detection performance, transparency in most machine learning models has been a major issue. Deep learning models tend to be black-box systems, making it hard to interpret what happens internally during their decision-making. Regulatory bodies in the financial world demand transparency in decision-making to ensure fairness, accountability, and adherence to financial rules. Explainable AI has thus become a necessary research direction as it seeks to enhance transparency in AI-based financial systems [14].

Explainable AI methods provide insight into how machine learning models work and how they produce predictions in an interpretable manner. These methods allow analysts to gain insights into the causal factors underlying the identification of suspicious transactions in financial fraud detectors. Some of the articles clarify the significance of explainability in a financial decision-making setting, where interpretability becomes required to facilitate regulatory compliance and trust operations [15-17]. Moreover, it is suggested that integrated systems incorporating explainable AI and machine learning frameworks can be used to increase transparency and reliability in financial analytics [18].

The new trends have also investigated a distributed, collaborative framework for fraud detection based on federated learning and graph-based structures. Federated graph Learning frameworks enable banks to collaboratively detect fraud without transmitting sensitive data across companies, thereby maintaining privacy and enhancing detection [19]. Similarly, explainable federated learning models are suggested to enhance the transparency and confidence in the AI-powered banking systems [20]. These solutions question the significance of integrating AI, graph analytics, and explainability to address financial fraud and systemic risk.

It is also applicable to financial network analytics for identifying systemic financial risk, as well as for fraud identification. Correlated financial systems are subject to so-called cascading failures when a shock propagates across networks. It is possible to identify strong nodes and structural vulnerabilities in the networks of financial transactions using graph-based learning models that enable regulators to track systemic risks [21].

Despite these developments, recent research has often focused on improving predictive performance and has not offered any useful information to financial decision-makers. Furthermore, a few studies use graph analytics with explainable AI to extract node-specific risk data from financial networks. There is an urgent need to address this gap to create strong, transparent fraud detection systems. So, the paper explores the potential of interpretable AI and graph analytics to infer knowledge from financial network data, detect fraud, and track systemic risks.

From a financial perspective, the spread of fraud across transaction networks can be viewed through the lens of financial contagion theory, in which criminal activities are disseminated among interlinked entities. Nodes with high centrality can serve as channels for systemic risk transmission, aligning with systemic risk propagation theories and anti-money laundering (AML) monitoring systems.

The study question that is being asked is the one below:

What would be the easiest way to derive node-based risk data from financial network data using interpretable artificial intelligence and graph analytics to enhance fraud detection, identify systemic risk, and facilitate transparent regulatory decision-making?

Our study has following contributions:

- This research introduces a regulatory-oriented, interpretable graph-analytic framework that enables node-level financial risk scoring by combining structural network features with precision, optimized ensemble learning, and explainable AI techniques.
- The proposed framework combines graph-based structural metrics with optimized ensemble learning models to achieve higher fraud detection performance.
- This study enhances transparency and regulatory usability by incorporating explainable AI methods and network risk visualization, enabling financial analysts to identify high-risk nodes, fraud clusters, and transaction bridge structures within complex financial networks.

The remainder of this paper is structured as follows:

Section 2 focuses on related work; Section 3 emphasizes the research methodology adopted in this study; and Section 4 presents the results, their discussion, comparisons, and the limitations of this study. Section 5 concludes by summarizing the main points and presenting research implications.

2. Literature Review

The increased sophistication of financial transaction systems has led to the use of artificial intelligence methods to detect fraud and financial risks. Financial institutions can use AI-based fraud detection systems to process large amounts of transaction data and identify suspicious behavior more effectively than rule-based monitoring systems do. Research has shown that machine learning algorithms can be useful for detecting patterns in fraudulent transactions by learning from historical financial information [3]. Moreover, the AI-fraud detection systems were demonstrated to enhance business efficiency and minimize losses related to fraudulent actions [2].

Decision trees, ensemble learning algorithms, and support vector machines are machine learning models that have been widely used in fraud detection. These methods examine the properties of transactions, their behavior patterns and characteristics of customer activities to determine suspicious activities. Nonetheless, as financial transactions become increasingly intertwined, traditional machine learning models tend to fail to capture the intricate connections between financial entities. Recent research thus highlights the necessity of network-based analytical schemes taking into account the relational structure of financial transactions [21].

Graph-based machine learning techniques have become a potent tool for analyzing financial transaction networks. GNNs can model relations among objects in financial systems and identify latent patterns of fraudulent activity. Cheng, Zou [5] explain that GNNs will offer an effective framework for analyzing relationship-based financial data and identifying fraud patterns within network structures. In the similar vein, Motie and Raahemi [6] further reveal that graph-based learning models can preserve the connectivity patterns of financial entities and thus are more effective at identifying fraud cases than conventional machine learning methods.

Recent studies have proposed a number of sophisticated GNN models for detecting financial fraud. For example, encoder-decoder GNNs have been designed to support networks of credit card transactions and to identify patterns of fraud using relational feature learning [9]. Another type of neural network is metapath-guided, which has been suggested to elicit heterogeneous relationships in financial networks and enhance the performance of fraud detection [22]. Moreover, adaptive GNNs have been used to analyze transaction patterns and detect suspicious activities in financial systems [7].

Temporal GNNs have also contributed to fraud detection by accounting for changes in transaction relationships in predictive models. Financial transactions form complex, interconnected networks in which accounts, customers, and institutions interact through dynamic relationships. These interactions encode valuable structural information that can

reveal hidden fraud patterns and mechanisms for systemic risk propagation. Correspondingly, it has been suggested to use causal temporal GNNs to identify causal links among financial transactions to predict fraud and forecast risks more effectively [12].

The other works focus on enhancing graph learning models by incorporating network-level trust and influence measures. For example, global confidence-based GNNs help incorporate trust information into financial networks, thereby improving fraud detection [13]. The graph convolutional network models have also been used to identify influential nodes and detect suspicious activities in financial transaction networks [8]. These directions show that the network-based approach to feature extraction is critical for financial fraud detection systems.

Another significant research direction is the explanation of artificial intelligence and financial fraud detection systems. Explainable AI methods enable understanding of how machine learning models reach their conclusions, allowing an analyst to interpret model predictions and know what goes into the results of the fraud detection process. According to Deliu and Moldovan [23], transparency and compliance with regulation are critical in AI-led financial systems, and explainability is the key. On the same note, Mohamed, Abdelqader [15] argue that a clearer explanation of AI frameworks fosters trust and accountability in the financial decision-making process.

A number of studies have identified explainable patterns in fraud detection models that provide interpretable insights into the analysis of financial transactions. Zhou, Li [17] suggest that the explanatory AI concept is user-centric and enables analysts to comprehend the reasons behind the fraud detection predictions. Equally, Tiwari, Jain [18] indicate that explainable AI can be implemented with machine learning frameworks to increase transparency and improve the reliability of fraud detection systems. There are other research works on optimization-based explainable AI in determining relevant features in fraud detection work [24].

Recent studies also explore privacy-preserving, distributed AI systems for detecting financial fraud. Federated graph learning models enable financial institutions to cooperatively train an anti-fraud model without exposing their sensitive transactional information [16]. Moreover, to promote transparency in collaborative financial analytics settings, explainable federated learning methods have been suggested [20]. The developments indicate the significance of privacy protection and interpretable AI in financial risk management systems.

In general, the available literature shows that there has been considerable advancement in the use of AI, graph analytics, and explainable AI for detecting financial fraud and analysing risk. Simultaneously, explainable AI methods increase transparency and support regulatory oversight. There has been scant research on mechanisms that integrate graph analytics with interpretable AI to provide node-level risk information to the financial decision-maker. The identification of this research gap is the basis for the current research.

3. Methodology

3.1. Research Framework

The proposed artificial intelligence framework in this paper reveals the content of fraudulent transactions and the systemic risk pattern using financial network data, based on a description of the financial network intelligence model that is intelligible to users. The methodology integrates learnable structural attributes via a graph, machine learning classification methods, imbalance control techniques, and explainable AI solutions to make accurate predictions and provide information to decision-makers in finance. The steps involved during the research process include adding the transaction network to the pre-processing and feature extraction systems, and training and optimizing the model based on different machine learning models. The Synthetic Minority Over-Sampling Technique (SMOTE) is also applied to address class imbalance, as is typically the case with fraud detection data during training [25]. Lastly, post-hoc explanation methods provide the necessary interpretability to enable regulators to understand the factors that drive the predicted risk of fraud at the global and transaction levels.

The Elliptic node-level risk indicator is defined as follows:

$$Risk(i) = \propto . \mathrm{P}_{\mathrm{illicit}}(i) + (1 - \propto). \mathrm{C(i)} \qquad (1)$$

Where $\mathrm{P}_{\mathrm{illicit}}(i)$ denotes the predicted probability of a transaction being illicit; C(i) represents the normalized structured centrality and $\propto$ is a weighting parameter, which controls the contribution of predictive and structural risk components.

A baseline model was proposed using Node2Vec to provide a graph-learning benchmark. Node2Vec conducts biased random walks on the transaction network and learns low-dimensional node representations that capture structural proximity and similarity in the neighbourhood. The resulting node representations served as input features to an

XGBoost classifier used to identify fraudulent transactions. The baseline will enable a comparison of conventional machine learning models that predict features of a handcrafted network with learning methods based on graph embedding. Figure 1 is the illustration of research flow diagram.

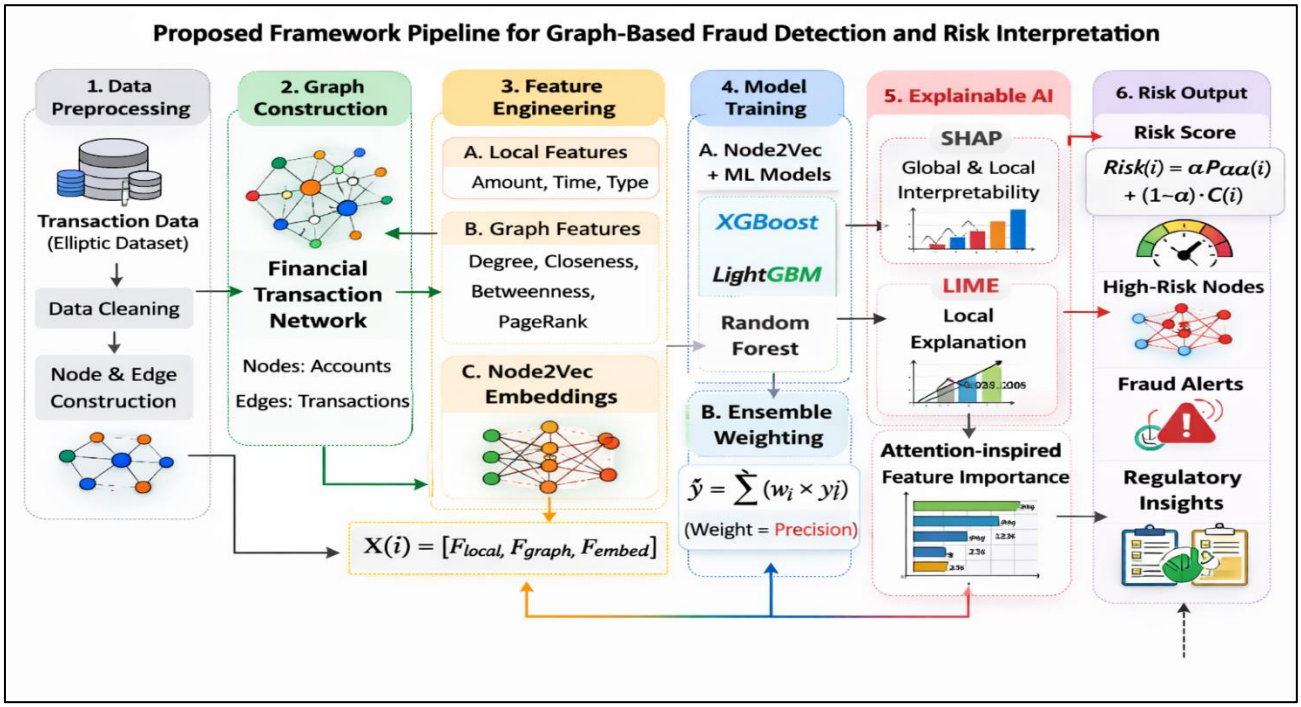

Figure 1 Research flow diagram

3.2. Dataset Construction and Feature Engineering

This study proposes to employ a dataset with a large scale transaction network. The transactions are labelled either licit or illicit, which provide learning to models for fraud detection. The fraudulent transaction constitute a small portion of dataset and class imbalance may hinder the correct classification of illicit transactions. To overcome this problem, SMOTE is used to generate synthetic examples of instances from the minority classes during training. SMOTE works by extrapolating new minority samples along the line segments connecting the existing minority observations in feature space. The method is better than simply duplicating samples (to balance class distributions) to improve model learning, as it fixes overfitting and generally improves generalization performance. The balanced training data enables machine learning models to learn more discriminatory behaviors related to fraudulent activities.

In addition to the initial transactional attributes, graph-based structural features were calculated to capture the relational aspects of the financial network. These characteristics comprise centrality metrics in the network, such as betweenness centrality and closeness centrality, which quantify the centrality of transactions within the network. High centrality values. Trades that serve as intermediate or bridging nodes of money flow networks can be a sign of suspicious activity, including money laundering chains or organized fraud rings.

3.3. Machine Learning Models

In the proposed framework, several machine learning models were implemented and compared to assess the effectiveness of various learning paradigms. These models are LightGBM (Optimized), Precision-Weighted Ensemble, XGBoost (Cost-Sensitive), XGBoost (Tuned), Random Forest (Calibrated), Gradient Boosting (Tuned), and Logistic Regression (Balanced).

XGBoost was fitted in two forms: the cost-sensitive form, which imposes a larger penalty for misclassifying illicit transactions, and the tuned model, specified via a hyperparameter search. The models trained included Random Forest and Gradient Boosting, which provide strong ensemble baselines that capture complex feature interactions. Class balancing was implemented using logistic regression to determine the advantages of nonlinear models in fraud pattern recognition.

A precision-weighted ensemble model was built to further improve predictive performance. It is based on the combination of products of several base classifiers. This ensemble method combines model predictions, assigning higher weight to classifiers with higher precision on the validation data. Maximizing false positives has been a priority in such weighting because, in regulatory fraud detection systems, it is necessary to reduce as many false positives as possible, since compliance teams can otherwise be overwhelmed by the overall illusion of compliance. This stage involves training and validating the ensemble model using the weighting strategies, which improve predictive performance and reduce false positives. The ensemble prediction is determined as follows:

$$\hat{y} = \sum w_i \, x \, y_i \qquad (2)$$

Where y_i represents the prediction of i_th model and w_i is its weight. A precision weighted criterion prioritizes models with the lower false positive rates.

3.4. Model Training and Evaluation

A supervised learning framework was used to train the machine learning models by splitting the training and testing data temporally to create a realistic experience of financial monitoring. We employed cross-validation method to optimize the model parameters, such as tree depth, learning rate and regularization factors through hyperparameter tuning. To evaluate the performance of models, classification performance metrics, including Area Under the Receiver Operating Characteristic Curve (AUC), Precision, Recall, and F1 score are used in this study. Although AUC measures the classifier's overall performance in detecting fraud, precision and recall describe the model's ability to accurately detect fraudulent transactions while minimizing false detections. The fact that the precision-recall curve also provides insight into the trade-off between detecting illicit transactions and controlling false positives. A confusion matrix analysis was also used to evaluate the true-positive, false-positive, and false-negative predictions of both models.

3.5. Interpretable AI Techniques

Interpretable AI techniques were also introduced into the suggested framework to increase transparency and regulatory usability. Post-hoc explanation methods such as SHAP (Shapley Additive Explanations) and LIME (Local Interpretable Model-Agnostic Explanations) were applied to explain model predictions and identify the most important features affecting the risk of fraud.

Moreover, an attention-inspired analysis is implemented using SHAP values to determine feature importance for individual predictions. This technique helps identify the dominant structural features that influence fraud risk and provides an interpretable approximation of the attention mechanisms used in deep learning models. Although traditional attention mechanisms are commonly associated with neural networks, the paper utilizes attention-like analysis to highlight the significant structural characteristics of the financial network.

3.6. Innovativeness of the Proposed Approach

The originality of the methodology lies in combining graph-based structural analysis, ensemble machine learning, and interpretable AI techniques into a single financial fraud detection model. This method, unlike the traditional fraud-detection system, uses only a transactional attribute, while network topology is used to identify the relational patterns hidden in the flow of financial transactions. Moreover, the application of precision-weighted ensemble modelling also addresses the issue of false alerts related to regulatory settings. The combination of explainable AI tools like SHAP and LIME enables the company to see the details of its predictions and thus provides financial institutions and regulators with a better understanding of what drives the risk of fraud. As a result, the suggested methodology offers a scalable, interpretable solution to the problem of detecting illicit financial operations and aids in data-driven management of financial risks.

4. Results and Discussion

4.1. Dataset and Experimental Setup

The Elliptic Bitcoin Dataset consists of 203,769 Bitcoin transactions across 49 temporal time steps, each with 166 features (94 local transaction-level and 72 one-hop neighborhood aggregates), and was used to experiment on 234,354 directed flow edges [26]. This dataset is accessible from the source [27]. Of all the 46,564 labelled transactions, 4,545 (9.76) are illicit, 90.24% are licit, and 157,205 transactions have not been labelled. The 166 Elliptic features were further augmented with eight graph-structural features of betweenness centrality, PageRank, clustering coefficient, core

number, in/out-degree, degree ratio and total degree to create a 174-dimensional feature of per-node. Five-fold stratified cross-validation was used to evaluate stability, with a typical temporal train/test split (train: t 34 or less; test: t 35 or greater) applied to avoid temporal leakage.

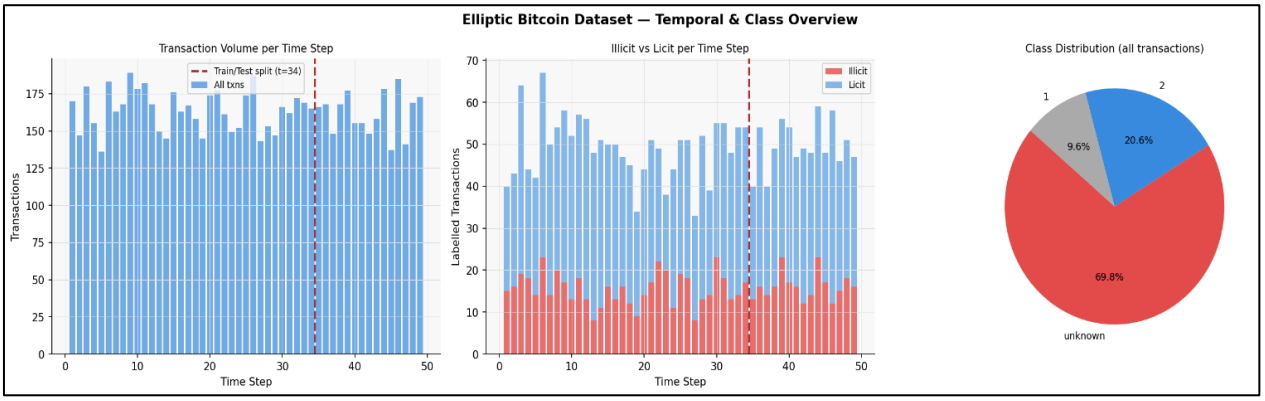

Figure 2 Elliptic dataset: volume of transactions per time (left); illicit and licit distribution when train/test cut-off at t =34 (center); class distributions in general (right). The temporal clustering of illicit activity indicates the organized criminal campaigns and promotes the temporal split protocol.

4.2. Model Performance

Table 1 Models performance comparison summary

Model	AUC	PR-AUC	Precision	Recall	F1 Score	Threshold	FPR	TP	FP
LightGBM (Optimized)	0.9383	0.8110	0.9750	0.7303	0.8351	0.80	0.0013	195	5
Precision-Weighted Ensemble	0.9397	0.8093	0.9604	0.7266	0.8273	0.72	0.0000	0	0
XGBoost (Cost-Sensitive)	0.9450	0.8100	0.8778	0.7266	0.7951	0.50	0.0000	0	0
XGBoost (Tuned)	0.9323	0.8033	0.7787	0.7378	0.7577	0.50	0.0000	0	0
Random Forest (Calibrated)	0.9221	0.7566	0.7059	0.7191	0.7124	0.52	0.0205	192	80
Gradient Boosting (Tuned)	0.9413	0.7922	0.6748	0.7228	0.6980	0.80	0.0239	193	93
Logistic Regression (Balanced)	0.8682	0.2944	0.1382	0.9251	0.2405	0.88	0.3953	247	1540
Node2Vec +XGBoost	0.9518	0.8163	0.8178	0.7378	0.77757	0.40	0.0280	799	178

Table 1 shows that tree-based ensemble models have the best results of fraud detection on the Elliptic Bitcoin dataset. The NodeVec + XGBoost combination has the highest AUC (0.9518) and PR-AUC (0.8163), indicating the best overall discrimination between illicit and legitimate transactions. Based on these results, the baseline model (Node2Vec + XGBoost) shows better capability to capture graph structural patterns than purely feature-based techniques. However, it shows lower precision and a higher false-positive rate than LightGBM, indicating a trade-off: improved graph representation increases detection power but slightly reduces decision reliability. Moreover, LightGBM (optimized) has the highest precision (0.9750) and the fewest false positives (5), indicating it is highly reliable for regulatory monitoring. The precision-weighted ensemble is also very precise, with no false positives, which is an advantage for reducing unnecessary investigations. Conversely, Logistic Regression has a high recall (0.9251) but produces too many false positives (1540) and cannot be practically deployed. On the whole, the ensemble boosting models offer the most cost-effective trade-off between precision, recall, and operational efficiency.

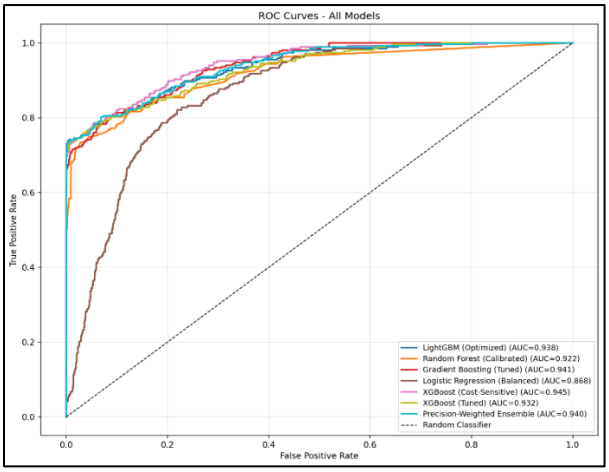

Figure 3 ROC curves of AI models

Figure 3 presents the ROC curves, which show that ensemble boosting models exhibit better discriminative performance in detecting fraud in the Elliptic Bitcoin transaction network. The XGBoost (cost-sensitive) model has the highest AUC (0.945); next are the precision-weighted ensemble (0.940) and Gradient Boosting (0.941), both showing good ability to differentiate between illicit and legitimate transactions at different thresholds. LightGBM (AUC = 0.938) is also quite competitive, with low false-positive rates and high true-positive rates, which are essential for minimizing unnecessary regulatory inquiries. In comparison, Logistic Regression (AUC = 0.868) demonstrates significantly worse performance, indicating a weak ability to accommodate more complex nonlinear transactional patterns. Notably, boosting models are interpretable using feature importance and explainability algorithms, enabling analysts to understand significant transaction patterns that affect fraud tendencies and to make clearer regulatory decisions.

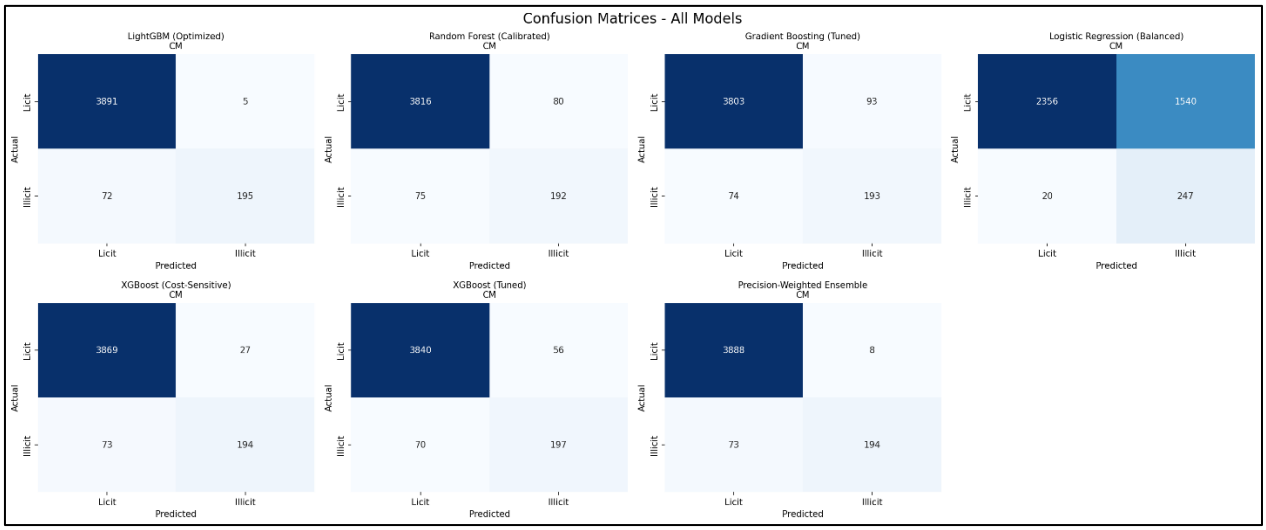

Figure 4 Confusion matrices of all seven supervised models

The confusion matrices shown in Figure 4 indicate that ensemble boosting models (LightGBM, XGBoost, and Gradient Boosting) achieve good performance in fraud detection, with high true positives and very low false positives. LightGBM, as well as the precision-weighted ensemble, is the least imbalanced, with the fewest misclassifications. Conversely, Logistic Regression produces very large numbers of false positives, making it unsuitable for financial fraud detection in highly imbalanced transaction datasets.

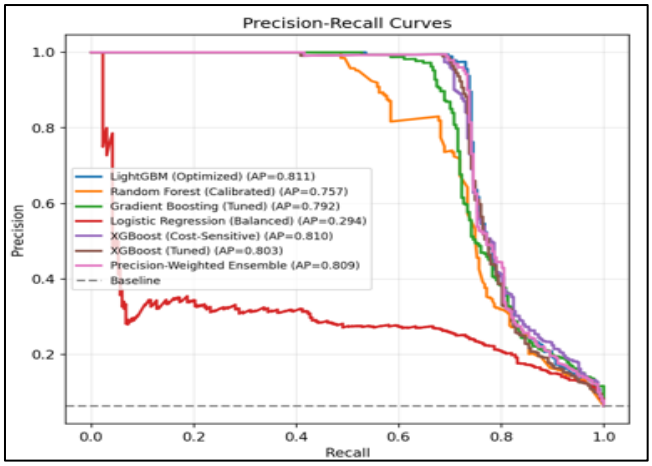

Figure 5 Precision-Recall curves obtained from implementation

Figure 5 presents the precision-recall curves, indicating that the ensemble boosting models achieve the best performance in fraud detection on the Elliptic Bitcoin dataset. At a large scale, LightGBM (AP = 0.811) and XGBoost (AP = 0.810) are precise at estimating recall and can be used to identify illegal transactions with few false recalls. The precision-weighted ensemble as a whole likewise operates in a comparable way, thereby explaining its strength against skewed financial data. Quite contrary, both gradient boosting and random forests portray a moderate decline in accuracy with a rise in recall. The results of a Logistic Regression (AP = 0.294) are very low, indicating that it cannot capture non-linearities,, and the patterns of transactions are extremely complicated. Overall, boosting-based models offer the most plausible trade-off between operational efficiency and detection accuracy in financial fraud monitoring systems.

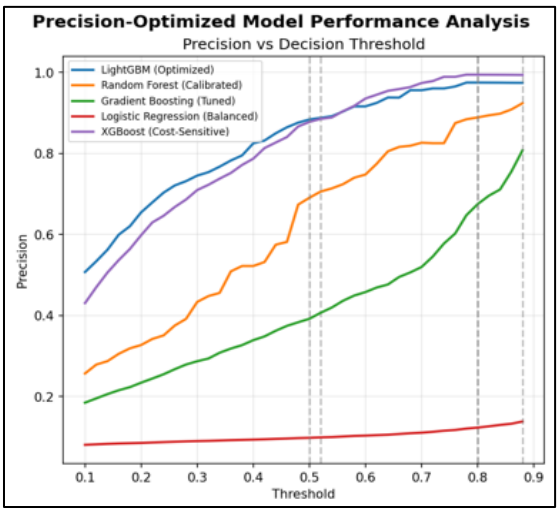

Figure 6 Variance in precision vs. decision threshold

Figure 6 shows the variation in precision across decision thresholds in the models. LightGBM and XGBoost are always the most precise, with values close to 1.0 at higher thresholds, indicating they are very effective at reducing false positives. Random Forest exhibits a moderate level of improvement with increasing thresholds, whereas Gradient Boosting exhibits a slow level of improvement, which represents a weaker stability of precision. Logistic Regression is extremely low at both levels, indicating poor applicability to imbalanced fraud detection problems. The discussion points out that boosting-based models offer a better option for precision control, enabling practitioners to use thresholds to balance detection quality and false-positive rates.

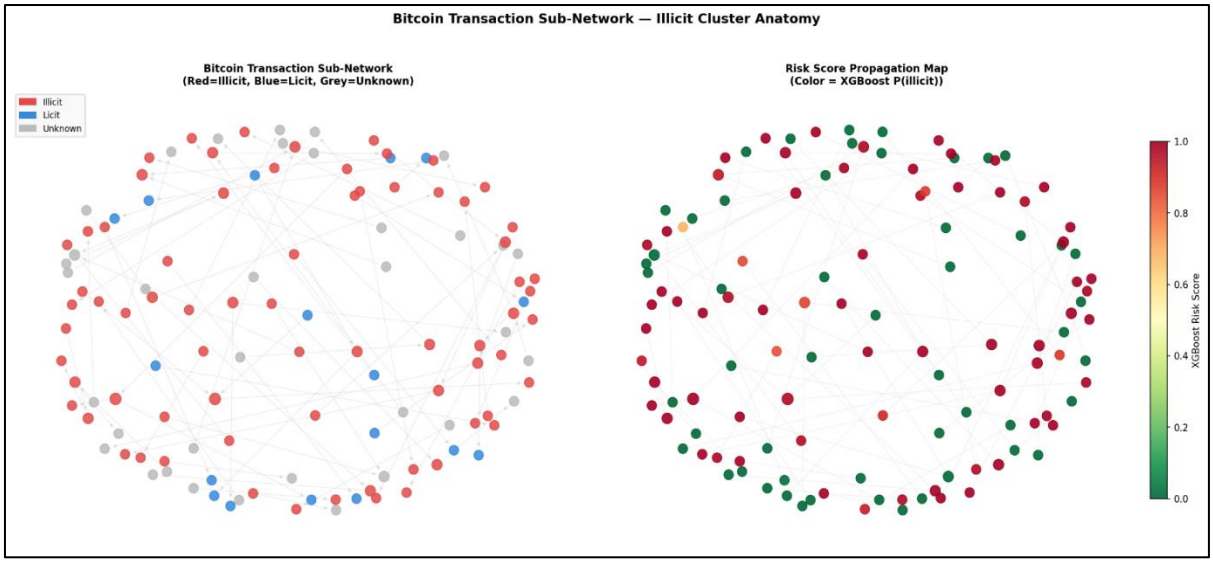

Figure 7 A Bitcoin sub-network of transactions.

Figure 7 shows a Bitcoin transaction sub-network, highlighting illicit clusters on the left side of the figure and corresponding node-level fraud risk scores predicted by XGBoost on the right side. This visualization illustrated the high-risk nodes from the interconnected patterns and propagated financial risk across the network

Table 2 Summary statistics of selected subgraph transaction features

Statistic	Tx_Id	True_label	Predicted_prob_gbm	Betweenness_centrality	Closeness_centrality
Count	5.000000e+00	5.000000	5.000000	5.0	5.000000
Mean	8.140942e+07	0.600000	0.238570	0.0	0.150000
Std	2.814236e+07	0.547723	0.379257	0.0	0.223607
Min	3.116838e+07	0.000000	0.003991	0.0	0.000000
25%	9.195272e+07	0.000000	0.007466	0.0	0.000000
50% (Median)	9.273526e+07	1.000000	0.125521	0.0	0.000000
75%	9.462359e+07	1.000000	0.149343	0.0	0.250000
Max	9.656717e+07	1.000000	0.906527	0.0	0.500000

Table 2 presents summary statistics for the subgraph features, providing information on the structural and predictive properties of the selected transactions in the financial network. The average of the true label variable (0.60) suggests that most nodes in the sample subgraph are illicit transaction nodes. The Predicted_prob_gbm values vary significantly (max = 0.9065), indicating that the Gradient Boosting model is efficient at distinguishing between low- and high-risk transactions. Although this small subgraph has zero values of betweenness centrality, the closeness centrality values (max = 0.50) indicate that some nodes play a relatively central role in the network. These results demonstrate the effectiveness of integrating graph-based features into predictive models to identify structurally influential transactions that may be related to illegal financial activity.

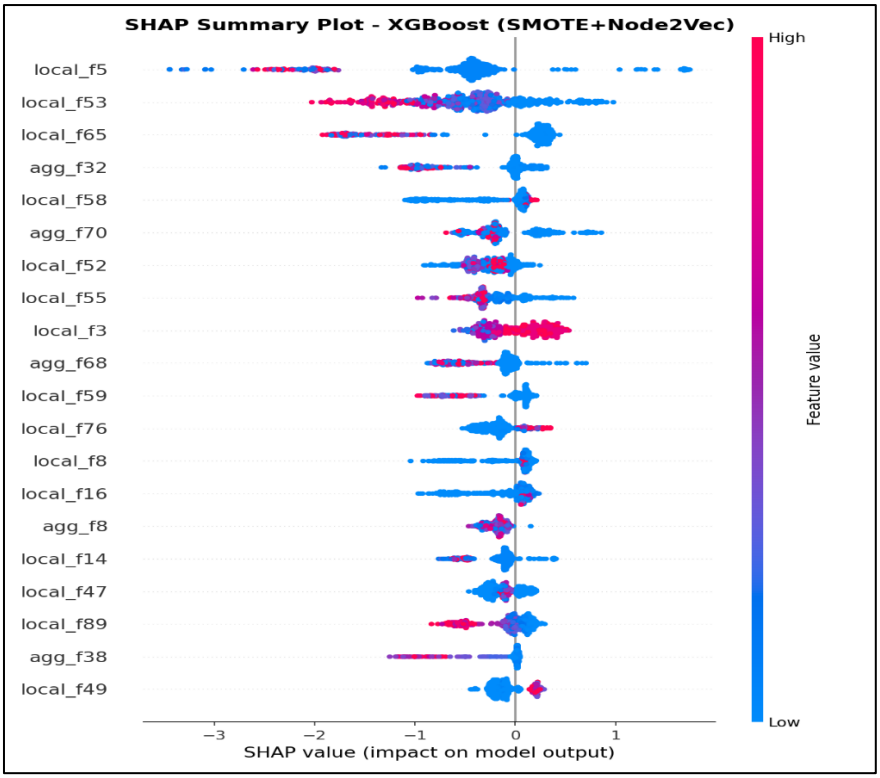

Figure 8 SHAP summary plot

Figure 8 illustrates that, using the SHAP summary plot, one can see that the features local f5, local f53, and local f65 have the highest impact on fraud prediction, and that the high- and low-impact features are clearly differentiated. The fact that SHAP values spread in a non-linear, asymmetric manner confirms that ensemble models such as XGBoost are appropriate. Meaningful risks can be identified by the presence of high feature value (red), which indicates the likelihood of fraud. This is relevant to our study because it shows that the model not only achieves high predictive accuracy but also provides feature-level and interpretable information that are crucial for knowledge extraction and regulatory decision-making.

4.3. Critical Discussion of Results

According to the experiment results, the ensemble-based and gradient-boosting models achieve high predictive accuracies in detecting fraud in a financial transaction network. In particular, LightGBM and cost-sensitive XGBoost achieve high AUC (0.938-0.945) and high precision-recall trade-offs, indicating they can help detect illicit transactions with a low number of false positives. The benefit of this precision-based structure is further enhanced in terms of regulatory relevance by reducing false alerts, which is one of the core elements of anti-money laundering (AML) and financial compliance machinery. The confusion matrix results show that the optimized LightGBM model is effective in detecting a massive percentage of illegal transactions at a low level of false positives. Furthermore, it can provide useful data on how transaction networks function, as centrality statistics of the graph can reveal the centrality of nodes, which in turn can identify influential nodes that can serve as facilitators during fraud cases. These findings demonstrate the feasibility of integrating graph analytics and interpretable machine learning to reveal more negative trends in financial crime and to detect risk factors that could be presented to regulators and financial institutions to explain.

Financially, nodes with high centrality can be viewed as potential centres of contagion, enabling the transfer of illegal financial transactions through the network. This aligns with AML compliance models, which emphasise detecting structurally influential nodes, as these are useful for early intervention and mitigating systemic risks.

4.4. Financial Risk Interpretation and Regulatory Implications

The research findings explain the use of predetermined financial risk models, specifically the theories of financial contagion and systemic risk propagation. Fraudulent activities can be transmitted through entities in an interconnected system of transactions similar to contagion, where risky nodes affect the transactions adjacent to them. Network centrality metrics, including betweenness and closeness centrality, identify highly structured nodes or intermediaries

that facilitate the flow of illicit funds. Table 2 reveals that nodes with higher predicted fraud probabilities are often associated with relatively higher closeness centrality, indicating their structural importance within the network. These nodes can influence the spread of illicit transactions by their central position in transaction flows. Although we observed limited betweenness centrality, closeness-based positioning remains a potential systematic exposure. In terms of compliance and AML, identifying these influential nodes can enable rapid detection of suspicious transaction chains, support targeted investigations, and facilitate regulatory monitoring by prioritizing risky participants in complex financial systems.

4.5. Comparison to Existing Studies

The findings of this paper are consistent with the latest study on the significance of network-based representations in financial fraud detection. For example, a distributed big data framework based on Node2Vec graph embedding and deep neural networks was proposed to learn structural patterns in financial transaction networks and shown to yield higher fraud detection accuracy through graph representation learning [28]. Similarly, Mao, Sun [29] came up with a knowledge-graph-based model, which uses related-party transaction relationships to complement fraud detection by taking into account inter-organizational dependencies. Wu, Chang [30] further built on this concept by developing an audit information knowledge graph to detect abnormal relationships among companies, auditors, and audit firms, helping regulators identify potential fraud entities using a path-based approach. Compared with these methods, the current work combines graph-based structural information with interpretable machine learning solutions that compete with them on predictive performance and also provide insight into both node-level risk and model interpretability, both of which are essential for regulatory decision-making and XAI in financial supervision. Our study demonstrates competitive or improved performance compared to recent studies by achieving higher precision, such as 0.9750, compared to a previous work with 0.9450 [31]. Another study reports ROC-AUC values of 0.8120 for wallets and 0.8920 for transactions using the Elliptic Bitcoin Dataset. In contrast, our study achieves higher performance, with AUC up to 0.9518 and PR-AUC exceeding 0.8100, indicating strong fraud detection capability under the evaluated dataset conditions and higher classification accuracy compared with the existing approach [32].

4.6. Limitations

The research has several methodological and evaluation weaknesses. First, graph characteristics like betweenness centrality are computationally intensive on large-scale networks; thus, an approximation based on sampling was used, which could introduce estimation variance, and on-the-fly transaction monitoring is not scalable. Second, LIME explanations are based on local linearization of a nonlinear model, which makes them vulnerable to the fidelity and stability of stochastic sampling. Third, the temporal train-test split assumes that transaction patterns are stationary, but model performance may degrade over time as concept drift occurs. Also, the research does not use a direct comparison of the GNNs baseline and does not provide formal regulation validation. Deep GNN models have shown strong relational learning abilities but tend to be less interpretable and more computationally expensive; hence, the embedded graph of Node2Vec provides a viable, understandable graph baseline for regulatory contexts. Lastly, the framework was tested in an offline batch environment and lacks support for real-time deployment, limitations on adversarial manipulation of graph features, and support for adversarial manipulation of graph features.

5. Conclusions

This paper presents a comprehensible artificial intelligence architecture for deriving knowledge from financial network data to aid in detecting fraud and risk. The framework successfully detects illicit transactions and concealed structural risks in financial networks by combining graph-based structural features and machine learning models. Experimental findings on the Elliptic Bitcoin dataset indicate that ensemble and boosting models, namely Node2Vec +XGBoost, LightGBM and cost-sensitive XGBoost, achieve high predictive accuracy with low false positives. The use of graph analytics enables the identification of nodes of influence and mediating transactions that may be favorable to fraud. Moreover, explainable AI methods enhance transparency and enable regulators in decision-making, as explainable risk insights can be interpreted by financial analysts and compliance enforcers.

It can be improved by future studies that introduce GNNs into this framework to more effectively represent complex relational dependencies in financial transaction networks. A combination of graph temporal learning and streaming analytics may enable fraud detection in real time and ongoing monitoring of financial systems.

References

[1] Yang, H., Z. Shukur, and S. Sahran, *A Review of Artificial Intelligence for Financial Fraud Detection.* Applied Sciences, 2026. **16**(4): p. 1931.

[2] Yaseen, H. and A.a. Al-Amarneh, *Adoption of artificial intelligence-driven fraud detection in banking: The role of trust, transparency, and fairness perception in financial institutions in the United Arab Emirates and Qatar.* Journal of Risk and Financial Management, 2025. **18**(4): p. 217.

[3] Pocher, N., et al., *Detecting anomalous cryptocurrency transactions: An AML/CFT application of machine learning-based forensics.* Electronic Markets, 2023. **33**(1): p. 37.

[4] Song, A., E. Seo, and H. Kim, *Anomaly VAE-transformer: A deep learning approach for anomaly detection in decentralized finance.* IEEE Access, 2023. **11**: p. 98115-98131.

[5] Cheng, D., et al., *Graph neural networks for financial fraud detection: a review.* Frontiers of Computer Science, 2025. **19**(9): p. 199609.

[6] Motie, S. and B. Raahemi, *Financial fraud detection using graph neural networks: A systematic review.* Expert Systems with Applications, 2024. **240**: p. 122156.

[7] Wang, W., et al. *AEFA: An Ensemble Framework for Fraud Detection in the Forex Market*. in *International Conference on Advanced Data Mining and Applications*. 2025. Springer.

[8] Islam, S., et al., *Detecting Fraudulent Transactions for Different Patterns in Financial Networks Using Layer Weigthed GCN.* Human-Centric Intelligent Systems, 2025. **5**(2): p. 181-195.

[9] Cherif, A., et al., *Encoder–decoder graph neural network for credit card fraud detection.* Journal of King Saud University-Computer and Information Sciences, 2024. **36**(3): p. 102003.

[10] Qian, J. and G. Tong, *Metapath-guided graph neural networks for financial fraud detection.* Computers and Electrical Engineering, 2025. **126**: p. 110428.

[11] Tian, Y. and G. Liu, *Spatial-temporal-aware graph transformer for transaction fraud detection.* IEEE Transactions on Industrial Informatics, 2024. **20**(11): p. 12659-12668.

[12] Duan, Y., et al., *Cat-gnn: Enhancing credit card fraud detection via causal temporal graph neural networks.* arXiv preprint arXiv:2402.14708, 2024.

[13] Liu, J., Y. Tian, and G. Liu, *Global confidence degree based graph neural network for financial fraud detection.* arXiv preprint arXiv:2407.17333, 2024.

[14] Adamu, A.I., et al., *A systematic literature review of advanced machine learning techniques in wireless body area networks: Application, challenges, and future directions.* IEEE Access, 2025. **13**: p. 194729-194778.

[15] Mohamed, A., K. Abdelqader, and K. Shaalan, *Explainable Artificial Intelligence: A systematic review of progress and challenges.* Intelligent Systems with Applications, 2025: p. 200595.

[16] Kapale, R., et al. *Explainable AI for fraud detection: enhancing transparency and trust in financial decision-making*. in *2024 2nd DMIHER International Conference on Artificial Intelligence in Healthcare, Education and Industry (IDICAIEI)*. 2024. IEEE.

[17] Zhou, Y., et al., *A user-centered explainable artificial intelligence approach for financial fraud detection.* Finance Research Letters, 2023. **58**: p. 104309.

[18] Tiwari, R., et al., *Artificial Intelligence for Fraud Detection: A Bibliometric Exploration*, in *Exploring AI Implications on Law, Governance, and Industry*. 2025, IGI Global Scientific Publishing. p. 203-216.

[19] Pan, Z., et al. *2SFGL: a simple and robust protocol for graph-based fraud detection*. in *2022 IEEE International Conference on Cloud Computing Technology and Science (CloudCom)*. 2022. IEEE.

[20] Aljunaid, S.K., et al., *Secure and transparent banking: Explainable AI-driven federated learning model for financial fraud detection.* Journal of Risk and Financial Management, 2025. **18**(4): p. 179.

[21] Li, R., et al., *Internet financial fraud detection based on graph learning.* IEEE Transactions on Computational Social Systems, 2022. **10**(3): p. 1394-1401.

[22] Li, E., et al., *Efficient relation-aware heterogeneous graph neural network for fraud detection.* World Wide Web, 2025. **28**(5): p. 55.

[23] Deliu, D. and A. Moldovan, *Fintech 5.0: From Shadows to Clarity—The Future of Transparency in Financial Organizations*, in *Transparency in FinTech: Exploring Technological Advances, Regulation, and Societal Impact with Case Studies*. 2025, Springer. p. 285-347.

[24] Hajiyev, H., et al., *An Explainable AI-based Fraud Detection System Using Recursive Feature Elimination and Waterwheel Plant Optimization for Financial Transactions.* Engineering, Technology & Applied Science Research, 2025. 15(5): p. 28114-28119.

[25] Ghaleb, F.A., et al., Ensemble synthesized minority oversampling-based generative adversarial networks and random forest algorithm for credit card fraud detection. IEEE Access, 2023. 11: p. 89694-89710.

[26] Weber, M., et al., Anti-money laundering in bitcoin: Experimenting with graph convolutional networks for financial forensics. arXiv preprint arXiv:1908.02591, 2019.

[27] MIT, elliptic-bitcoin-dataset, H. Face, Editor. 2019.

[28] Zhou, H., et al., Internet financial fraud detection based on a distributed big data approach with node2vec. IEEE access, 2021. 9: p. 43378-43386.

[29] Mao, X., et al., Financial fraud detection using the related-party transaction knowledge graph. Procedia Computer Science, 2022. 199: p. 733-740.

[30] Wu, H., et al., Financial fraud risk analysis based on audit information knowledge graph. Procedia Computer Science, 2022. 199: p. 780-787.

[31] Elmougy, Y. and L. Liu. Demystifying fraudulent transactions and illicit nodes in the bitcoin network for financial forensics. in Proceedings of the 29th ACM SIGKDD conference on knowledge discovery and data mining. 2023.

[32] Pérez-Cano, V. and F. Jurado, Fraud detection in cryptocurrency networks—An exploration using anomaly detection *and heterogeneous graph transformers.* Future Internet, 2025. **17**(1): p. 44.