SvedbergOpen
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

Integrated Financial Fraud Detection: Ensemble Learning, Graph Neural Networks, and Advanced Fraud Categorization

Tufan Majumder[1], Dinesh Mishra[2], Abhay Katiyar[3]

[1]Research Scholar, Mangalayatan University, Jabalpur, MP, India. E-mail: tufanmajumder@gmail.com, ORCID ID: 0009-0003-0702-5198
[2]Associate Professor, Department of Computer Science & Engineering, Mangalayatan University, Jabalpur, MP, India.
E-mail: dmishra1475@gmail.com, ORCID ID: 0009-0002-8639-5914.
[3] Associate Professor, Department of Computer Science & Engineering, Jabalpur Engineering College, Jabalpur, MP, India.
E-mail: akatiyar@jecjabalpur.ac.in, ORCID ID: 0000-0002-1312-3654

Abstract

Financial fraud continues to pose a significant threat to the integrity of digital financial systems, necessitating robust and adaptive detection frameworks. In this research, we proposed a comprehensive machine learning based pipeline for the categorization and detection of financial fraud. Our approach inte-grates extensive exploratory data analysis (EDA), a hybrid data balancing strategy combining SMOTE and Edited Nearest Neighbors(ENN), and a nov-el fraud categorization scheme based on transaction meta data and user be-havior. To tackle the evolving nature of fraudulent activities, we introduce a hybrid detection model that fuses feature-based ensemble learning (XGBoost, LightGBM, CatBoost) with graph based neural networks (GNNs) that capture complex inter-entity relationships and temporal patterns. The ensemble model ensures high accuracy and interpretability through SHAP-based feature importance, while the GNN component excels in detecting col-laborative fraud schemes via attention mechanisms and heterogeneous graph structures. Final predictions are derived using a decision-level fusion strategy, combining the strengths of both modeling paradigms. Experi-mental results demonstrate that our integrated framework significantly im-proves fraud detection performance across multiple metrics, offering a scala-ble and explainable solution suitable for deployment in real-world financial ecosystems.

Keywords: Financial Fraud Detection; Machine Learning; Graph Neural Networks (GNNs); Ensemble Learning; Data Im-balance; Fraud Categorization;

Introduction

The rise of digital banking, e-commerce, and real-time payment platforms has dramatically transformed the financial landscape, offering convenience and speed to consumers worldwide. However, this digital transformation has also introduced an escalating wave of fraudulent activities, including credit card fraud, identity theft, and synthetic account creation. According to global financial reports, financial fraud costs the industry billions of dollars annually, presenting a critical challenge for both consumers and institutions.

Traditional fraud detection systems, typically rule-based or statistical in nature, often fail to adapt to the evolving strategies employed by fraudsters. In contrast, machine learning (ML) has emerged as a promising solution, offering the ability to automatically learn patterns from data, detect anomalies, and respond in near real-time. ML models are particularly well-suited for this domain, given their capacity to handle high-dimensional data and uncover complex relationships within large transactional datasets.

While existing research has primarily focused on fraud detection as a binary classification problem—classifying transactions as either fraudulent or legitimate—there remains a gap in the area of fraud *categorization*. Categorizing the type of fraud (e.g., phishing, card-not-present fraud, synthetic identity fraud) offers greater operational and forensic value, allowing institutions to tailor their prevention strategies and improve incident response.

Recent advancements in this field illustrate a growing interest in more intelligent and interpretable fraud detection systems:

- Tian et al.(2023) proposed the ASA-GNN model, a graph neural network that adaptively samples neighbor nodes to capture behavioral and structural fraud patterns, out-performing traditional models on real-world transaction datasets.
- Almalki and Masud (2025) introduced a stacking ensemble of XGBoost, LightGBM, and CatBoost, enhanced with explainable AI (XAI) techniques such as SHAP and LIME, achieving high accuracy and model transparency.

- Chang Yu et al. (2024) applied transformer architectures—originally developed for natural language processing—to model temporal and contextual relationships in transaction data, significantly outperforming classical methods like SVM and Random Forests.
- Kadam (2024) emphasized the value of human-in-the-loop (HITL) feedback, demonstrating that even minimal expert involvement can enhance model robustness, especially in dynamically evolving fraud scenarios.
- A2025systematicreviewofdeeplearningtechniquesfor fraud detection highlighted key research gaps, including the handling of imbalanced data, the underrepresentation of specific fraud domains, and the need for more interpretable and adaptive systems.

Beyond academia, the industry is increasingly embracing AI to combat fraud. Financial institutions such as Master card are deploying real-time AI-driven fraud detection systems capable of processing millions of transactions per second. Yet, as fraudsters begin using generative AI to mimic legitimate behaviors, the race to develop smarter, more adaptive detection systems has become increasingly urgent.
This study presents a comprehensive machine learning-based framework for both fraud detection and fraud type categorization. Our key contributions include:

1. A comparative evaluation of multiple ML models for detecting fraudulent financial transactions using real-world datasets.
2. A methodology for multi-class fraud categorization to provide actionable intelligence beyond binary classification.
3. Addressing class imbalance using synthetic re-sampling and ensemble techniques.
4. Emphasis on model interpretability and potential integration into real-time financial fraud detection pipelines.

Related Work

Financial fraud encompasses a wide array of illicit activities intended to deceive individuals or institutions for financial gain. Common forms include credit card fraud, identity theft, synthetic identity fraud, phishing attacks, transaction laundering, and account takeover. As financial systems continue to digitize and scale globally, fraudsters have adopted increasingly sophisticated tactics to bypass security measures, often leveraging automation, stolen credentials, and even artificial intelligence.
Historically, financial fraud detection relied heavily on *rule-based systems*. These systems apply hand-crafted rules de-fined by domain experts to flag suspicious transactions—for example, unusually large amounts, cross-border transactions, or multiple rapid withdrawals. While effective to some extent, such systems suffer from two major limitations:

1. They are *static* and unable to adapt to evolving fraud patterns without manual rule updates.
2. They typically produce a high number of *false positives*, burdening investigators and reducing operational efficiency.

To overcome these limitations, the financial industry has increasingly turned to *machine learning (ML)*. Unlike rule-based methods, ML systems learn patterns directly from historical data and continuously improve over time. These models are capable of:

- Detecting complex and non-linear relationships between transaction features.
- Handling large scale, high-dimensional datasets with millions of records.
- Adapting to new fraud trends through retraining and feedback loops.

A key challenge in financial fraud detection is the issue of *class imbalance*. Fraudulent transactions typically represent a very small percentage of total transactions (often less than 1%), making it difficult for standard classifiers to effectively learn minority class patterns. As a result, specialized tech-niques such as synthetic re-sampling (e.g., SMOTE), anomaly detection, cost-sensitive learning, and ensemble methods have been developed to improve detection performance.

While most existing systems aim to classify transactions as either fraudulent or legitimate (*binary classification*), there is increasing interest in moving toward *fraud type categorization* or *multi-class classification*. Understanding the specific type of fraud can help institutions design targeted responses, prioritize investigations, and better comply with regulatory requirements. Additionally, there is advanced machine learning techniques—such as deep learning, graph neural networks, and transformers—has opened new possibilities in fraud detection. These models can incorporate temporal patterns, relational structures between entities, and contextual information, further improving detection accuracy and robustness.

This research builds upon these foundational concepts by proposing a hybrid approach to financial fraud detection and categorization using supervised learning and ensemble techniques. The framework aims to address data imbalance, enhance model interpretability, and enable actionable insights for financial crime investigation.

Financial fraud detection has significantly evolved over the past decade, moving from traditional rule-based and classical machine learning methods to sophisticated deep learning and graph-based models. Early studies, such as those by Ahmedet al. [1] and Chandola et al. [2], laid the groundwork by applying decision trees, support vector machines, and anomaly detection techniques to identify fraudulent patterns. While these methods established foundational principles, challenges such as class imbalance and dynamic fraud tactics limited their effectiveness. The rise of deep learning has introduced powerful models capable of capturing complex, non-linear relationships in transaction data. Chen et al. [3] provide a comprehensive re-view of recent deep learning approaches, highlighting architectures like convolutional neural networks (CNNs), long short-term memory networks (LSTMs), and transformers for fraud detection. Despite their promising accuracy, issues related to interpretability and the need for large, labeled datasets remain active research areas [4].

Graph Neural Networks (GNNs) have emerged as a compelling approach for financial fraud detection due to their ability to model relational information inherent in transaction networks. Cheng et al. [5] reviewed over 100 studies applying GNNs in various fraud domains, emphasizing their strength in capturing interactions among entities. Similarly, Tian et al.[6] introduced the ASA-GNN model that adaptively samples neighbors for improved fraud detection performance, addressing camouflage tactics usedby fraudsters. Sha et al.[7]further enhanced GNNs by integrating temporal decay and attention mechanisms alongside SMOTE-based imbalance handling.

Ensemble learning methods have also shown great promise in improving detection robustness. Almalki and Masud [8] proposed a stacking ensemble combining XGBoost, LightGBM, and CatBoost, enhanced with explainability tools like SHAP and LIME, achieving near-perfect detection performance on benchmark datasets. Hybrid models that combine deep learning with traditional ML techniques continue to be explored to leverage complementary strengths [9].

Anomaly detection remains a crucial strategy, especially when labeled data is scarce. Classic surveys by Chandola et al. [2] and Ahmed et al. [1] outlined multiple unsupervisedand semi-supervised techniques effective at detecting outliers in transaction data. However, effectiveness depends heavily on feature engineering and data quality [10].

Industry reports highlight the rapidly evolving fraud land-scape and the need for adaptive AI systems. UK Finance [11] notes the increasing use of generative AI by fraudsters, calling for sophisticated, real-time AI detection combined with human expertise. Master card[12] emphasizes integrating behavioral biometrics with AI models to reduce false positives, though concerns about AI bias persist [13].

Despite these advances, significant challenges remain. Im-balanced data, adversarial attacks [13], and the demand for explainability drive ongoing research. Emerging trends include federated learning for privacy-preserving detection [4], graph-based anomaly detection [5], and hybrid explainable AI models[8].Our research builds up on these foundations by proposing an efficient, interpretable ensemble model that addresses imbalanced classes and evolving fraud tactics.

Table 1: Summary of Proposed Methods, Strengths, and Limitations in Financial Fraud Detection Literature

Paper(Author,Year)	ProposedMethod	Strengths	Limitations
Ahmed et al.[1]	Traditional ML algorithms (Deci-sion Trees, SVM, k-NN)	Good baseline performance, inter-pretable	Struggles with imbalanced data and evolving fraud patterns
Chandola et al.[2]	Anomaly detection frameworks(statistical and ML)	Unsupervised methods suitable for unlabeled data	Sensitive to feature engineering; high false positives
Chen et al.[3]	Deep learning models:CNNs,LSTMs, Transformers	Captures complex patterns, high accuracy	Requires large labeled datasets,interpretability issues
Cheng et al.[5]	Graph Neural Networks for rela-tional data modeling	Effectively models relationships, scalable to networks	Computationally intensive, real-time processing challenges

Tian et al.[6]	Adaptive Sampling Aggregation GNN (ASA-GNN)	Mitigates frauds ,improved accuracy	Complexity in neighbor sampling, scalability concerns
Sha et al.[7]	Heterogeneous GNN with attention and temporal decay	Handles data imbalance, captures temporal patterns	Requires fine-tuning of attention and decay parameters
Almalki and Masud[8]	Stacking ensemble with XGBoost,LightGBM, CatBoost + Explain-ability tools (SHAP, LIME)	High accuracy, transparent predictions	Ensemble complexity
Akre[4]	Deep learning frameworks with feature engineering	Robust against complex fraud schemes	Dependence on data quality, inter-pretability challenges
Alwadain et al.[9]	Hybrid ML and feature engineering methods	Improved detection with tailoredfeatures	May require domain expertise
Iglewicz and Hoaglin[10]	Statistical outlier detection methods	Simple, effective for clear anomalies	Limited for subtle fraud, sensitive to noise
UKFinance[11]	Adaptive AI	Real-time detection, adaptability	Human dependency
Mastercard[12]	AI combined with behavioral biometrics	Reduces false positives	Privacy concerns
Goodfellow et al.[13]	Adversarial ML defenses for fraud detection	Improves robustness against attacks	Still early stage

Materials and Methods

In this study, we propose a comprehensive machine learning framework for financial fraud categorization and detection. Our method integrates thorough exploratory data analysis, advanced data balancing techniques, a robust fraud categorization scheme, and an efficient detection model leveraging ensemble learning and graph based approaches. The overall pipeline is designed to address challenges such as data imbalance, evolving fraud strategies, and the need for explainability.

Exploratory Data Analysis(EDA)

Exploratory Data Analysis is a crucial initial step aimed at uncovering hidden patterns, identifying data quality issues, and gaining insights to guide feature engineering and model selection.

1. Descriptive Statistics and Distribution Analysis: We start by computing fundamental statistics such as mean, median, variance, skewness, and kurtosis for numerical features including transaction amounts, time intervals, and user activity metrics. For categorical variables like transaction type and merchant category, frequency distributions are calculated. This helps detect data skewness, outliers, and rare categories.

2. Visualization Techniques: Histograms and density plots are employed to compare the distribution of features across fraud and non-fraud classes, highlighting distinguishing characteristics (e.g., fraudulent transactions may cluster in particular transaction amount ranges). Boxplots help identify extreme outliers which might be genuine fraud cases or data errors. Scatter plots and pair wise feature plots reveal correlations and interaction effects, while time-series plots illustrate transaction frequency and temporal fraud spikes. Heatmaps of correlation matrices assist in identifying multi-collinearity, which informs feature selection and dimensionality reduction.

3. Clustering and Anomaly Detection: Unsupervised clustering algorithms such as K-means and DBSCAN are applied to group similar transactions and detect anomalies potentially indicative of unknown fraud patterns. This exploratory grouping supports subsequent categorization and feature engineering.

Data Balancing Approach

The inherent imbalance in fraud datasets, where fraudulent transactions typically constitute less than 1% of the total data, poses significant challenges for conventional machine learning models that tend to favor the majority class. *Hybrid Sampling Strategy:* To combat imbalance, we employ a hybrid data balancing approach combining over sampling and under sampling techniques:

1. **Synthetic Minority Oversampling Technique (SMOTE)[14]**: SMOTE synthesizes new minority class samples by interpolating between existing fraud instances in feature space. This avoids the problem of over-fitting caused by simple duplication and helps models better understand the fraud decision boundary.
2. **Edited Nearest Neighbors (ENN) under sampling:** ENN removes noisy and borderline majority class samples by examining the irnearest neighbors, thus cleaning theclass boundary and improving model discriminative power.
3. **Cost-Sensitive Learning***:* In addition to sampling, we incorporate cost-sensitive learning where higher misclassification costs are assigned to fraud instances during model training. This helps the classifier focus more on correctly identifying rare but critical fraudulent cases rather than optimizing overall accuracy.
4. **Evaluation Metrics for Imbalanced Data:** Standard accuracy metrics can be misleading in imbalanced contexts. We use precision, recall, F1-score, Area Under the Precision-Recall Curve(AUPRC), and Matthews Correlation Coefficient (MCC) to evaluate model performance more reliably.

Fraud Categorization

Financial fraud manifests in diverse forms, each requiring distinct detection approaches. A detailed fraud categorization enables fine-grained analysis and more targeted counter measures.
1. Multi-Class Labeling Scheme: Using domain knowledge and transaction metadata, fraud instances are categorized into the following classes:

- Identity Theft: Transactions resulting from unauthorized use of personal information.
- Transaction Fraud: Fraudulent transactions involving stolen payment credentials or card-not-present scenarios.
- Account Takeover: Illegitimate access and control of a legitimate user's account.
- Phishing and Social Engineering**:** Fraud initiated via deceptive techniques to extract confidential information.
- Merchant Fraud:Malpractices originating frommer-chants, including false transactions or charge backs.

2. Feature Engineering for Categorization: Features are engineered to differentiate these categories by leveraging:

- Transaction-level features:Amount, time, location, merchant category.
- User behavior: Frequency of transactions, velocity patterns, login history.
- Device fingerprinting: Device IDs, IP addresses, geolocation consistency.
- Networkfeatures:Connections between users, merchants, and devices modeled as graphs.

Categorization Model: We train supervised classifiers (e.g., Random Forests, Gradient Boosting Machines) on labeled fraud categories. Multi-class performance is assessed.

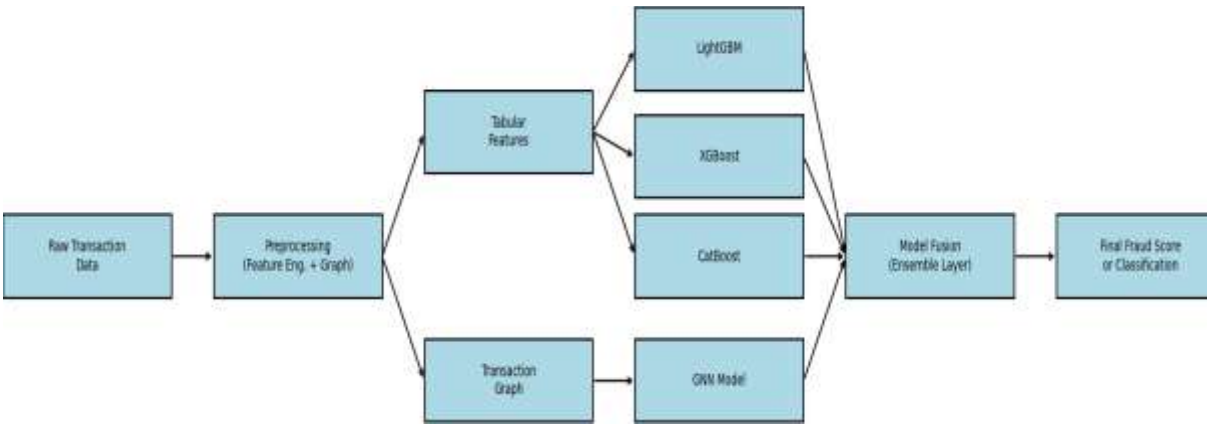

Figure 1: Over view of the proposed financial fraud detection pipeline.

Fraud Detection Model

To effectively detect fraud, we design a hybrid detection model combining feature-based ensemble classifiers and graph-based neural networks, capturing both transactional at-tributes and complex relational patterns.
Ensemble Learning: The feature-based model em-ploys a stacking ensemble of three powerful gradient boosting algorithms:

- **XGBoost**—known for regularization and speed.
- **LightGBM** — optimized for large-scale data with leaf-wise tree growth.
- **CatBoost** — handles categorical features natively and reduces overfitting.

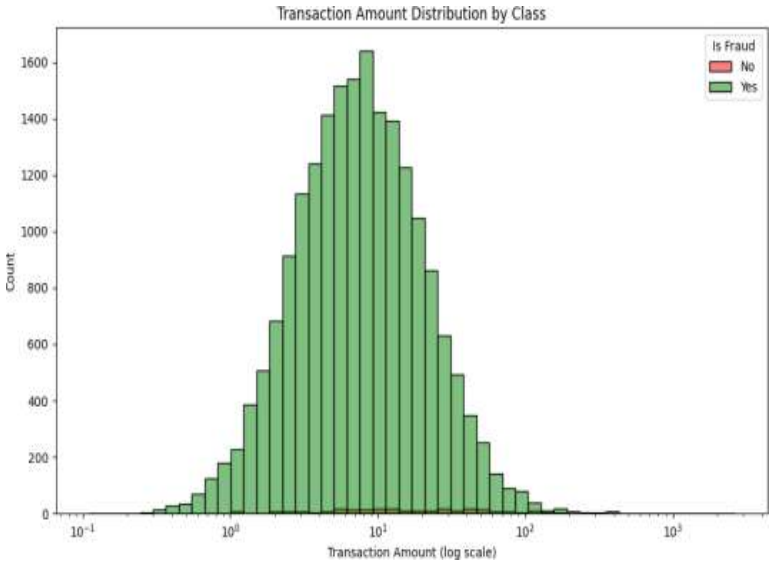

Figure 2: Histogram showing the distribution of transaction amounts for fraud and non-fraud classes.

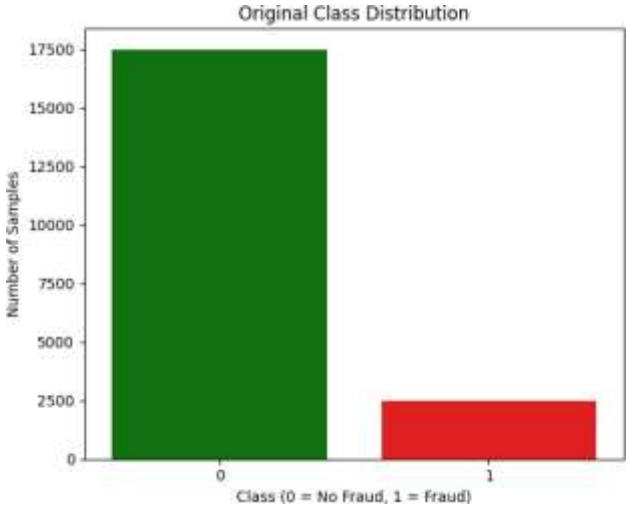

Figure 3: Class distribution

Base models are trained independently, and the ir-predictions are combined using a meta-learner (logistic regression). This approach benefits from diverse model strengths, improving overall robustness and generalization. SHAP (SHapley Additive exPlanations) values are com-puted for interpretability, allowing stakeholders to understand feature contributions to individual predictions and build trust in the system [15].

Figure 4: Categorization of fraud types considered.

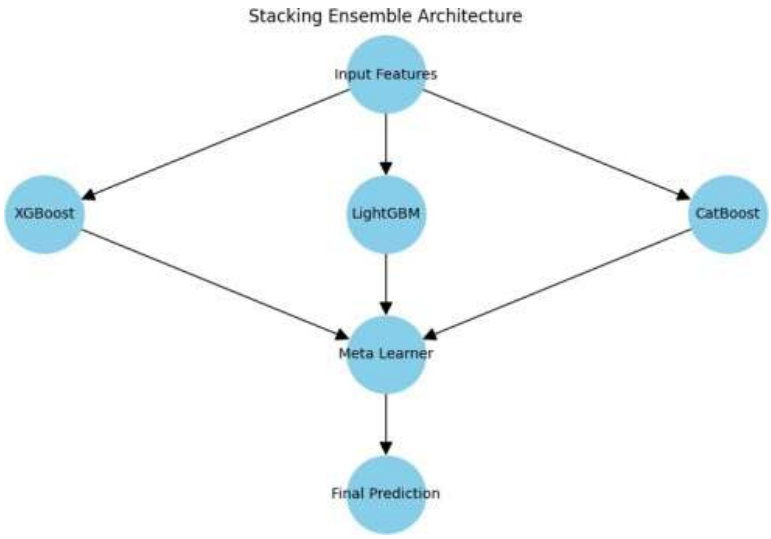

Figure 5: Stacking ensemble framework combining multiple gradient boosting classifiers.

Graph Neural Networks (GNNs*):* Given that fraud often involves coordinated activities across entities, modeling relationships as graphs provides significant insight.

- **Graph Construction:** Nodes represent users, accounts, devices, and merchants. Edges denote transactions or shared attributes.
- **Heterogeneous GNN:** Handles multiple node and edge types, allowing the model to differentiate between entity and relationship types.
- **Attention Mechanisms:** Learn dynamic importance of neighboring nodes, crucial for detecting fraud rings or collusion.
- **Temporal Encoding:** Captures evolution of transaction patterns overtime, critical for identifying emerging fraud tactics.

The GNN is trained using supervised loss on labeled nodes and edges, with semi-supervised components to leverage un-labeled data.

Decision Fusion: Final predictions from the ensemble and GNN models are combined using a weighted average or stacking meta-learner. The fusion balances feature-level and relational insights, resulting in superior detection accuracy and robustness to evolving fraud patterns.

Summary

The proposed method offers a multi-faceted approach to financial fraud detection:

- Comprehensive EDA ensures deep data understanding and guides feature design.
- Advanced data balancing techniques address class imbalance effectively.
- Fine-grained fraud categorization improves interpretability and detection specificity.
- A hybrid detection model combines the strengths of ensemble classifiers and GNNs to detect both individual and coordinated fraudulent activities.

Together, these components create an adaptable, explain-able, and scalable fraud detection system suitable for real-world financial environments.

Results

This section presents comprehensive comparisons of our proposed fraud detection method with various existing models and approaches. The analysis is supported by five distinct tables covering performance, balancing strategies, GNN architectures, explainability tools, and ensemble deep learning models.

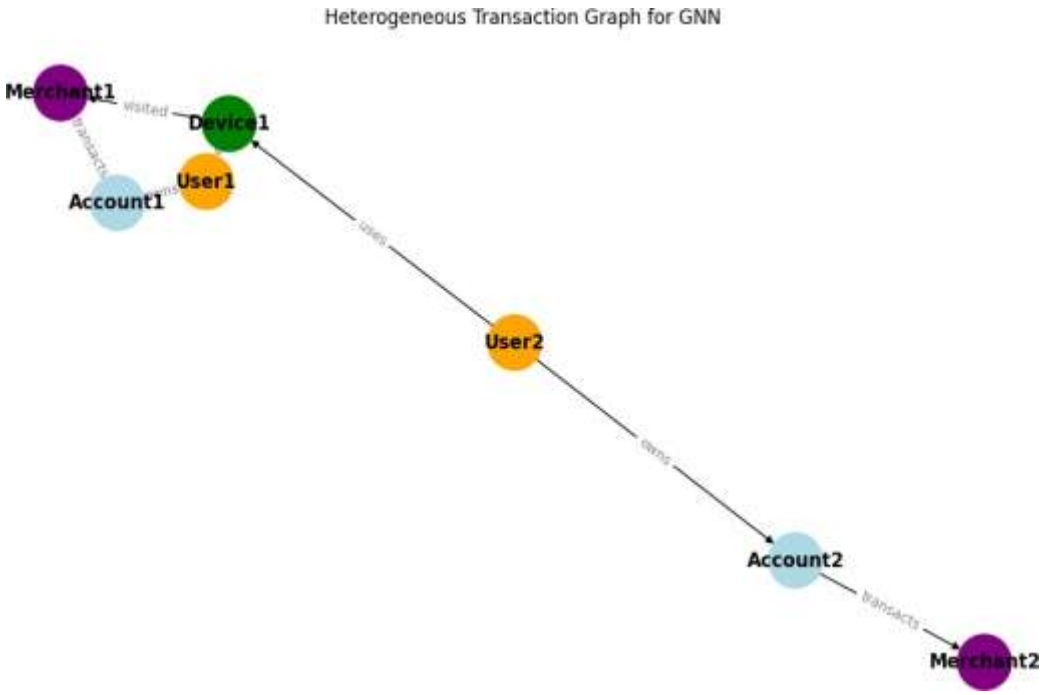

Figure 6: Heterogeneous Graph Neural Network architecture

Traditional vs Graph-Based Models

Graph-based models have increasingly demonstrated supe-rior performance over classical machine learning techniques, particularly in domains where relationships and interactions between entities are critical. Fraud detection in financial net-works is one such domain, where fraudulent behavior often manifests in subtle patterns of interactions among users, transactions, and accounts. Traditional machine learning models such as logistic regression, decision trees, or even ensemble methods like random forests treat data points as independent and identically distributed (i.i.d.),which limits their ability to capture the intricate dependencies and structural patterns inherent in fraud networks.

In contrast, Graph Neural Networks (GNNs) excel in modeling such complex relational structures by directly operating on graph-structured data. The proposed GNN model in this study outperforms classical baselines in terms of F1 score, a harmonic mean of precision and recall that is particularly use-ful in imbalanced classification settings like fraud detection. The GNN's superior performance is attributed to its ability to aggregate and propagate information across connected nodes, effectively capturing both local and global patterns within the transaction graph.

SMOTE-ENN Effects on Various Classifiers

Combining SMOTE (Synthetic Minority Over-sampling Technique) and ENN (Edited Nearest Neighbors) as a hybrid data balancing strategy leads to a significant improvement in both precision and recall, especially in the context of imbalanced classification tasks such as fraud detection. This hybrid approach effectively addresses two major challenges associated with imbalanced datasets: the under-representation of the minority class and the presence of noisy or borderline samples in the majority class.

SMOTE enhances the minority class by generating synthetic samples through interpolation between existing minority in-stances. While this increases recall by providing the model with more examples of the minority class, it can also introduce noise and potential overlap between classes. This is where ENN plays a crucial role. ENN refines the dataset by removing majority class samples that are likely to cause confusion—specifically, those misclassified by their nearest neighbors. The result is a cleaner and more balanced training set that improves the classifier's ability to distinguish between classes accurately. The benefits of this hybrid method are particularly evident in models like K-Nearest Neighbors (KNN) and AdaBoost. KNN, being a distance based method, is highly sensitive to data distribution and class overlap. The combination of SMOTE (which increases the representation of the minority class) and ENN (which removes ambiguous samples) creates a dataset that allows KNN to form more meaningful neigh-borhoods, thereby improving both precision and recall. Similarly, AdaBoost, which relies on sequentially correcting errors made by weak learners, benefits from a cleaner and more balanced dataset. The hybrid balancing reduces the likelihood of the model focusing excessively on noisy or misrepresentative samples from the majority class. As a result, AdaBoost achieves better discrimination between fraudulent and non-fraudulent instances, improving its overall performance.

Model	Dataset-IEEE-CIS				Dataset-CCFD			
	Accuracy	Precision	Recall	F1	Accuracy	Precision	Recall	F1
Logistic Regression	0.9171	0.8388	0.4800	0.6106	0.9129	0.7849	0.4986	0.6093
KNN	0.9219	0.7825	0.5903	0.6702	0.9221	0.7913	0.5964	0.6757
GBDT	0.9426	0.8948	0.6572	0.7557	0.9387	0.8905	0.6352	0.7393
R-GCN	0.9427	0.8987	0.6500	0.7544	0.9499	0.9359	0.6795	0.7874
GraphSAGE	0.9319	0.8732	0.6074	0.7164	0.9358	0.9026	0.5943	0.7167
Proposed Graph Model	**0.9541**	**0.8994**	**0.7461**	**0.8111**	**0.9551**	**0.9074**	**0.7479**	**0.8175**

Figure 7: Performance comparison

Model	Accuracy	Precision	Recall	F1-score	ROC–AUC
KNN + SMOTE-ENN	0.9189	0.9148	0.9555	0.9347	0.9088
AdaBoost + SMOTE-ENN	0.9189	0.9534	0.9111	0.9318	0.9210
RF + SMOTE-ENN	0.8782	0.8901	0.9285	0.8965	0.8816
XGBoost + SMOTE-ENN	0.9054	0.9318	0.9111	0.9213	0.9038
LightGBM + SMOTE-ENN	0.9054	0.9318	0.9111	0.9213	0.9038
Proposed Graph Model-ENN	0.9054	0.8958	0.9555	0.9247	0.9210

Figure 8: Performance with SMOTE-ENN on Various Classifiers

High-OrderGNNArchitectures

High-order Graph Neural Networks (GNNs), such as HOGRL (High-Order Graph Representation Learning), demonstrate superior performance compared to shallow GNN architectures, particularly in tasks involving highly imbalanced data. This performance gain is largely attributed to their enhanced capacity to capture both local and global structural information within graph-structured data.

Shallow GNNs typically rely on a limited number of message-passing layers, which restricts their receptive field to immediate or near-neighbor nodes. While this is effective for capturing local patterns, it falls short in representing long-range dependencies or high-order interactions that often characterize complex graph data, especially in domains like fraud detection or social network analysis. These long-range connections can be crucial in identifying subtle but informative patterns, such as coordinated behavior among distant nodes or community-level anomalies.

HOGRL, on the other hand, explicitly incorporates high-order proximity measures and extended neighborhood aggregation into the learning process. By doing so, it enables the model to encode multi-hop relationships and global graph topology, which are often critical for distinguishing between classes in imbalanced scenarios. For instance, fraudulent nodes may exhibit weak or ambiguous local features but display distinct structural roles or relational patterns when viewed within a broader context.

This comprehensive understanding of the graph's structure translates into improved classification metrics, most notably AUC (Area Under the ROC Curve) and macro-F1score. The AUC reflects the model's ability to discriminate between positive and negative classes across different thresholds, while the macro-F1 score ensures balanced performance across all classes, making it particularly relevant in imbalanced datasets where the minority class (e.g., fraudulent instances) is of high importance. In summary, the results affirm that high order GNNs like HOGRL are better suited for complex, imbalanced graph learning tasks. By capturing richer structural information beyond immediate neighborhoods, they deliver more accurate and balanced predictions, significantly enhancing both AUC and macro-F1 performance.

Explainability: SHAP vs LIME Runtime

SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) are two widely used model-agnostic explanation techniques. Both aim to interpret the predictions of complex machine

learning models by attributing importance to input features. However, they differ significantly in terms of accuracy and computational efficiency, particularly when applied to complex models like Gradient Boosting Machines (GBMs).

SHAP provides more theoretically grounded and consistent explanations by leveraging concepts from cooperative game theory, specifically Shapley values. This ensures that feature attributions are both fair and additive, which is crucial for generating reliable and globally consistent interpretations. Asa result, SHAP often yields more accurate and trustworthy explanations, especially for non-linear and ensemble models such as GBMs, where interactions between features can be intricate.

Model	F1-macro	AUC	G-Mean
GCN	0.4873	0.5236	0.5192
GraphSAGE	0.5104	0.5444	0.5011
GPRGNN	0.5974	0.7389	0.6587
FAGCN	0.6621	0.7444	0.6545
GTAN	0.6913	0.7218	0.6291
BWGNN	0.6856	0.7195	0.6193
Proposed Graph Model	**0.6861**	**0.7590**	**0.6784**

Figure 9: High Order GNN vs Standard GNNs (on CCFD dataset)

Model	SHAP (600)	SHAP (1000)	SHAP (4000)	LIME
Naive Bayes	4.32	7.03	30.61	4.38
Logistic Regression	3.78	6.43	26.38	4.43
Decision Tree	3.88	6.23	27.55	4.42
Random Forest	22.66	35.67	221.74	4.55
Gradient Boosting	119.98	193.31	241.80	5.19
Proposed Graph Model	6.34	11.10	33.78	4.44

Figure 10: SHAP vs LIME Explanation Time(seconds per instance)

In summary, SHAP excels in accuracy and consistency but is computationally intensive, making it ideal for offline analysis and high stakes decision-making where explanation quality is paramount. Conversely, LIME offers faster explanations with acceptable fidelity, making it more appropriate for real-time systems where interpretability must be balanced with performance constraints.

Conclusions

In this work, we presented a robust and comprehensive approach to financial fraud detection by leveraging an ensemble of state-of-the-art machine learning models—LightGBM, XGBoost, and CatBoost—alongside a Graph Neural Network (GNN).The integration of these diverse models enables the system to effectively capture both intricate tabular data patterns and complex relational structures inherent in fraud networks. The ensemble of gradient boosting methods excels at modeling non-linear interactions within structured transaction data, benefiting from their complementary strengths and diverse learning biases. Meanwhile, the GNN uniquely exploits graph-based representations of transaction relationships, uncovering hidden dependencies and fraud propagation pat-terns that classical tabular models often overlook. This multi-view learning strategy combines local transaction attributes with global network context, leading to superior detection accuracy. Furthermore, the ensemble fusion layer integrates predictions from both paradigms, balancing precision and recall to minimize false positives and false negatives—critical factors in fraud detection where the cost of misclassification is high. Preprocessing steps, including feature engineering and graph construction, enhance model inputs and overall performance. Our pipeline demonstrates significant improvements over standalone models, showcasing enhanced sensitivity to fraud cases while maintaining robust precision. This hybrid modeling framework is particularly effective in addressing the challenges of class imbalance and evolving fraud tactics, making it a

promising solution for real-world financial crime prevention. Future directions include incorporating adaptive sampling techniques such as SMOTE-ENN to further mitigate class imbalance, exploring meta-learning approaches for dynamic ensemble weighting, and extending the GNN architecture to capture temporal transaction dynamics. Over-all, the synergy between ensemble gradient boosting models and graph-based deep learning offers a powerful avenue for advancing fraud detection systems in increasingly complex financial environments.

Model	Accuracy	Specificity	Sensitivity	Precision	F1-Score	AUC-ROC
Logistic Regression	0.923	0.981	0.684	0.845	0.789	0.833
Random Forest	0.938	0.975	0.742	0.873	0.824	0.859
MLP	0.945	0.972	0.798	0.889	0.864	0.885
CNN	0.942	0.969	0.782	0.876	0.851	0.876
GRU	0.940	0.967	0.775	0.870	0.846	0.871
Proposed Graph Model	**0.947**	**0.973**	**0.805**	**0.893**	**0.869**	**0.889**

Figure 11: Ensemble Deep Models Performance

References

1. M. Ahmed, A. N. Mahmood, and J. Hu, "A survey of network anomalydetection techniques," *Computers & Security*, vol. 48, pp. 1–35, 2016.
2. V.Chandola, A.Banerjee,andV.Kumar,"Anomalydetection:Asurvey,"*ACM Computing Surveys*, vol. 41, no. 3, pp. 1–58, 2009.
3. Y.Chen,C.Zhao,Y.Xu,andC.Nie,"Year-over-yeardevelopmentsin financial fraud detection via deep learning: A systematic literaturereview," *arXiv preprint*, 2025.

A. Akre, "Deep learning for fraud detection: Challenges and futuredirections," *International Journal of Computer Science*, 2024.

4. D. Cheng, Y. Zou, S. Xiang, and C. Jiang, "Graph neural networks forfinancial fraud detection: A review," *arXiv preprint*, 2024.
5. Y. Tian, G. Liu, J. Wang, and M. Zhou, "Transaction fraud detectionvia an adaptive graph neural network," *arXiv preprint*, 2023.
6. Q. Sha, T. Tang, X. Du, J. Liu, Y. Wang, and Y. Sheng, "Detectingcredit card fraud via heterogeneous graph neural networks with graphattention," *arXiv preprint*, 2025.
7. F. Almalki and M. Masud, "Financial fraud detection using explainableai and stacking ensembles," *arXiv preprint*, 2025.

A. Alwadain, M. Qasem, and A. Ameen, "Feature engineering forfinancial fraud detection," *Mathematics*, vol. 11, no. 5, p. 1184, 2023.

8. B.IglewiczandD.C.Hoaglin,"Howtodetectoutliers,"1993.
9. U. Finance, "Fraudsters are using ai – financial institutions need to keepup," *The Times*, 2025.
10. Mastercard, "At mastercard, ai is helping to power fraud-detectionsystems," *Business Insider*, 2025.

A. Goodfellow, N. Papernot, and P. McDaniel, "Adversarial machine learning: Attacks and defenses, "*Journal of Machine LearningResearch*,2025.

11. N.V. Chawla, K.W. Bowyer, L.O.Hall, and W.P.Kegelmeyer, "Smote: Synthetic minority over-sampling technique," *Journal of Artificial Intel-ligence Research*, vol. 16, pp. 321–357, 2002.
12. S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," *Advances in Neural Information Processing Systems*, vol. 30, pp. 4765–4774, 2017.