Enhanced Integrated Model for Financial Fraud Detection Using Graph Machine Learning

S.Babu[1*], V.Rama Narayanan[2], T.Nirmal Raj[3], J.Srinivasan[4]

[1,2,3,4]Sri Chandrasekharendra Saraswathi Viswa Mahavidyalaya, Kanchipuram, Tamil Nadu, India.

babulingaa@ac.in*, vgram31@hotmail.com,
tnirmalraj@kanchiuniv.ac.in, jsrinivasan@kanchiuniv.ac.in

Abstract. Fraud in Financial domain bearings a major hazard to financial systems of modern era, causing in significant commercial losses and lessen the trust among the stakeholders. Conventional fraud detection techniques based on traditional algorithms of machine learning habitually fail to identify the compound relational patterns present in large-scale financial transaction data. In order to deal this limitation, this study proposes an intelligent graph-based machine learning framework for financial fraud detection. By modeling financial entities and their interactions as graphs, the proposed approach leverages Graph Machine Learning (GML) techniques to uncover hidden relationships, structural patterns, and anomalous behaviors associated with fraudulent activities. Furthermore, the impact of data imbalance—a common challenge in fraud datasets—is analyzed, and appropriate balancing strategies are applied to enhance detection performance. Experimental results exhibit that the graph-based approach significantly outstrips the conventional methods in identifying fraudulent transactions, highlighting its effectiveness in improving accuracy, robustness, and reliability in real-world financial fraud detection systems.

Keywords: Graph-Based Fraud Detection, Financial Transaction Networks, Graph Machine Learning, Data Imbalance Handling

1 Introduction

The rapid digitalization of financial services has commanded to an exponential growth in the size and difficulty of financial transactions. Though this transformation has improved operational efficiency and customer convenience, it has also created new opportunities for fraudulent activities. Financial fraud has consequently become a major concern for financial institutions, regulatory authorities, and consumers, resulting in significant economic losses and erosion of trust in financial systems [1]. Detecting fraudulent transactions in a timely and accurate manner remains a challenging task due to the dynamic and adaptive nature of fraud schemes.

Conventional machine learning techniques have been widely adopted for fraud detection due to its facility to analyze huge datasets and recognize suspicious patterns. However, these approaches exhibit several limitations. First, traditional ML models heavily rely on large volumes of labeled data for effective training. In real-world financial datasets, fraudulent instances are relatively rare, leading to highly imbalanced data that degrades model performance. Second, many advanced ML methods, predominantly deep learning-based methods, function as black-box systems, suggesting limited interpretability. This lack of transparency makes it difficult for stakeholders to understand, validate, and trust the detection outcomes.

R. Vasanth Kumar Mehta et al. (eds.), *Proceedings of the International Conference on Intelligent Systems for a Sustainable Future (ISSF 2026)*, Atlantis Highlights in Intelligent Systems 16,
https://doi.org/10.2991/978-94-6239-693-7_115

Graph Machine Learning (GML) has emerged as a powerful alternative to address the shortcomings of traditional approaches. By representing financial transactions, users, accounts, and devices as nodes and their interactions as edges, graph-based models capture the underlying relational structure of financial data. This explicit modeling of relationships enables GML techniques to uncover complex dependencies and hidden patterns that are often overlooked by conventional ML methods [1].

Unlike traditional models, graph-based approaches can effectively utilize relational information, allowing accurate fraud detection even with limited labeled data. Furthermore, the structural representation of graphs enhances model interpretability, as fraudulent behavior can be traced through observable connections and interaction patterns. By analyzing transactional networks and behavioral relationships, GML models significantly improve the identification of anomalous and coordinated fraud activities.

2 Literature Review

Recent research aims to optimize graph database algorithms for faster traversal, hybridize graph and relational models, and enhance distributed graph computing for scalability. Researchers are integrating graph databases with AI and ML for predictive analytics. This approach is particularly useful in fraud detection, supply chain management, and cyber security.

Soroor Motie., et.al, (2024) focused on several relevant studies, which reveals categorizing GNN methodologies into models relevant to Convolutional, Attention, metapath, and transformer based [1]. The review suggests that supervised and semi supervised learning are the main focus of most of the applications. Very limited focus on unsupervised methods. Furthermore, the review detects the gap in edge and graph level detection of anomaly in the financial sector. The authors recommend future research directions to improve the efficiency of the GNN methods in contesting financial fraud.

Ayushi Patil., et.al, (2024) reveals the use of model related to graph based machine learning to detect fraudulent transactions [2]. It powers Neo4j for graph modeling, for detecting real time anomaly by identifying the complex relationships present in the transaction data. The study also addresses the imbalance issue of the data using the Synthetic Minority Oversampling TEchnique (SMOTE) to further enhance accuracy of the model by balancing fraud and non-fraud cases in the dataset. Performance done based on the evaluation metrics like accuracy, recall and precision which confirm the efficiency of the model in improving fraud detection in banking systems.

Aam Albone., et.al, (2024) proposed the model that integrates machine learning techniques and graph databases to enhance the accuracy of fraud detection [3]. The hybrid approach aims to provide an additional level of security just by focusing on the relationship that exists between fraudsters or fraudulent cases. The results recommend that the graph based learning approach improves the fraud detection through the complex relationships that exists among entities. Which in turn results more reliable and consistent classification output when compared with the standalone machine learning models. In addition, the usage of graphs enhances the interpretability of machine learning models, which in turn makes the detection process more transparent.

Babu, S., et.al, (2017) recommended an effective technique called EMOTE for oversampling to solve the imbalance issues on the dataset [4]. In the proposed method, "each misclassified instance along with their nearest neighbor which was retrieved from correctly classified instances was populated in the actual dataset" [4]. The datasets with different imbalance ratio was considered for performance evaluation and also different experiments were executed. The results evidence that the classifiers were capable to enhance their performance on the dataset which was balanced by EMOTE.

3 Methodology

The proposed methodology, as shown in Fig. 1, starts by load the Bank Sim dataset into a Neo4J graph database to define the labeled property graph. The dataset was retrieved from Kaggle. In the defined graph, customers and merchants are the entities which was represented by nodes, the transactions happened between them was represented by edges and the transactions details are stored as properties. Graph databases are intended to reveal connected information, by considering the nodes to represent data points and the edges to establish the relationships between the nodes. By influencing the advanced querying techniques and the graph algorithms of Neo4J, the system was capable to reveal an outstanding result for the real time queries and anomaly detection.

As the next step, Data preprocessing was done which involves selection of apt features, normalizing the numerical data and encoding categorical variables are done. Since the dataset was imbalanced, EMOTE- Enhanced Minority Oversampling TEchnique was applied to balance the dataset before applying the machine learning techniques. The EMOTE confirms that balanced distribution was made on fraud and non fraud instances before the model training.

Once the preprocessing and balancing the dataset was over, graph based features was extracted by using the algorithms and functions present in Neo4j. The hidden interactions in between the customers and merchants were captured by deriving the structural and relational features like node degree, transaction frequency, pattern of connectivity and relationship on multi edges. For an instance, multiple transactions which happens on the same customer and merchant specifies that there was suspicious activity. The model based on graph improves the representation of transactional behavior and strengthen the ability of the model in detecting the anomalies.

After the dataset is updated by transactional and graph based features, the machine learning models like K-Nearest Neighbor, Support Vector Machine and XGBoost are applied to classify the transactions. Every model reveals its own patterns from the balance dataset to differentiate between fraudulent and non fraudulent transactions. The standard measures like Accuracy, Precision, Recall, F1 Score, ROC and AUC are used to evaluate the performance of the models. On the basis of comparative analysis, the most efficient model is considered for the final prediction of fraud or non fraud transaction. This proposed integrated approach that combine graph based relational analysis with the machine learning techniques was defined in order to improve the efficiency of fraud detection.

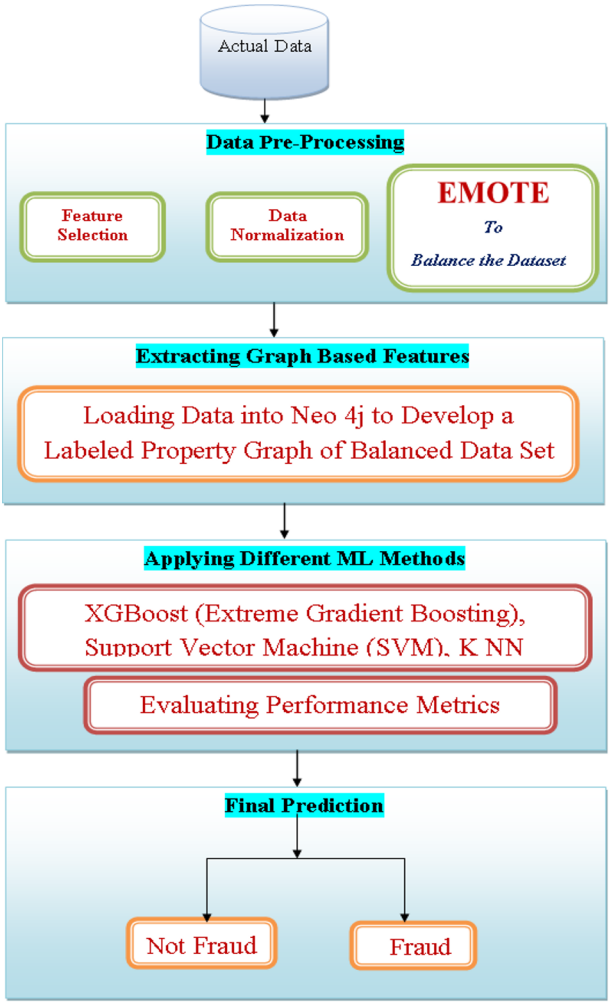

Fig. 1. Proposed Flow of Methodology

4 Experiments and Results

To evaluate the performance of the proposed model, Bank Sim dataset was used. It is a dataset which consists of simulated bank transaction created using a masked transactional data from the Spanish bank. The main purpose of the Bank Sim dataset is to generate synthetic data for fraud detection and also proved that it is closely related to the real world transactions. The dataset has totally 5,94,643 records, out of which 7200 records related to fraudulent transactions and 5,87,443 are the valid transactions.

The Bank Sim dataset consists of randomized simulated data; the values were differed from the actual financial transaction data. But it replicates key transactional patterns in an effective manner. The main advantage of this particular dataset is that, it does not contain null and duplicate values which in turn ensure the data consistency for the model training. Even then, the important challenge is that imbalance is present

in the fraudulent transactions. That is the non fraudulent transactions are about 81.59 times of larger than the fraudulent class. To ensure the accurate prediction of fraud detection, more careful selection of the training data is importance which will help the model to identify the effective fraudulent patterns exists on the dataset. Sensitivity, Specificity, Gmean, F-Measure and ROC are the Measures [5][6] used to evaluate the performance of the proposed model.

Accuracy

$$ACC = TP + TN / (TP + TN + FP + FN) \quad (1)$$

Sensitivity (True positive rate)

$$SN = TP / (TP + FN) \quad (2)$$

Specificity (True negative rate)

$$SP = TN / (TN + FP) \quad (3)$$

G-Mean (A Geometric mean of precision and recall)

$$G\text{-mean} = \sqrt{\boldsymbol{Precision\ X\ Recall}} \quad (4)$$

F-Measure (A harmonic mean of precision and recall)

$$F\text{-Measure} = 2\ X\ \boldsymbol{Precision\ X\ Recall} / (\boldsymbol{Precision}+\boldsymbol{Recall}) \quad (5)$$

To assess the performance of the proposed model, the following experiments were carried out and the outcomes are competed with other existing methods. They are

1. Performance of different classifiers on the imbalanced Bank Sim dataset and on the Dataset that balanced by the Technique named EMOTE.
2. Performance Comparison of proposed model on Bank Sim Dataset

Table 1. Characteristics of Data Sets.

Data Set Name	Total Number of				Imbalance Ratio
	Attributes	Instances	Minority Class Instances	Majority Class Instances	
Bank Sim	10	5,94,643	7200	5,87,443	81.59

4.1 Evaluation of EMOTE Using Various Classifiers

To evaluate the technique, EMOTE, different classifiers was performed on the actual imbalanced data set. The outcomes were recorded and different measures were calculated. The EMOTE technique was applied and the data set was balanced. The classifiers are again applied on the balanced dataset. From the outcomes of the classifier, different measures were calculated. The Table 2 presents the values of calculated measures from both executions.

The values presented in the Table 2 contrast the performance of three classfiers namely SVM, KNN and XGBoost on the imbalance dataset and the dataset balanced by the EMOTE technique. In the imbalanced dataset classifiers are able to achieve lesser sensitivity because the classfiers favor the majority class. However, specificity is higher

because the classifiers are able to perform well on the dominant class. After the execution of the balancing method EMOTE, sensitivity increases significantly which reveals that better prediction on minority class instances. On the other hand, the overall accuracy also improves noticeably as shown in the Lifts by column of the table which exhibit the effectiveness of balancing the dataset.

Table 2. Performance of Various Classifiers on the Data Set Balanced by EMOTE

Classifiers	Sensitivity (%)		Specificity (%)		Overall Accuracy (%)		Lift By
	Actual Dataset	**Balanced by EMOTE**	Actual Dataset	**Balanced Dataset by EMOTE**	Actual Dataset	**Balanced Dataset by EMOTE**	
SVM	62.5	**56.39**	74.45	**88.44**	68.475	**87.42**	18.94
KNN	66.7	**33.33**	78.72	**56.27**	72.695	**84.8**	12.11
XG Boost	63.89	**90.28**	70.2	**91.68**	67.045	**90.98**	23.94

The other measures like G-Mean and AUC [7] were also gained a remarkable improvement that indicates, the model was able to perform well in both the classes on the balanced dataset. In addition, F1 score also improved based on the rise of recall, that proves the model was more effective in handling minority class predictions. Among the considered classifiers, XGBoost presents the highest improvement in overall accuracy, which suggest that the classifier benefits more from balanced dataset. The table. 3 results highlight the significance of dataset balancing technique EMOTE, particularly in the situations where predicting the minority class is vital such as fraud detection, churn prediction or medical diagnosis.

Table 3. G-Mean and AUC Measures of Classifiers on the Data Set Balanced by EMOTE

Classifiers	G-Mean		F1 Score		AUC	
	Actual Dataset	**Balanced Dataset By EMOTE**	Actual Dataset	**Balanced Dataset By EMOTE**	Actual Dataset	**Balanced Dataset By EMOTE**
SVM	0.609	**0.862**	0.586	**0.824**	0.6848	**0.823**
KNN	0.641	**0.841**	0.562	**0.812**	0.7269	**0.854**
XG Boost	0.772	**0.903**	0.714	**0.896**	0.6705	**0.912**

4.2 Evaluation of Proposed Model Using Various Classifiers

To further prove the efficiency of the proposed model, the dataset that was balanced by EMOTE is uploaded into NEO4j to change into a graph model with entities as nodes and relationships as edges. The same is illustrated in Fig. 2. The connections were established in between the entities based on relationship that exists on them. There after machine learning methods were applied on the graph database and the results were recorded and shown in table 4.

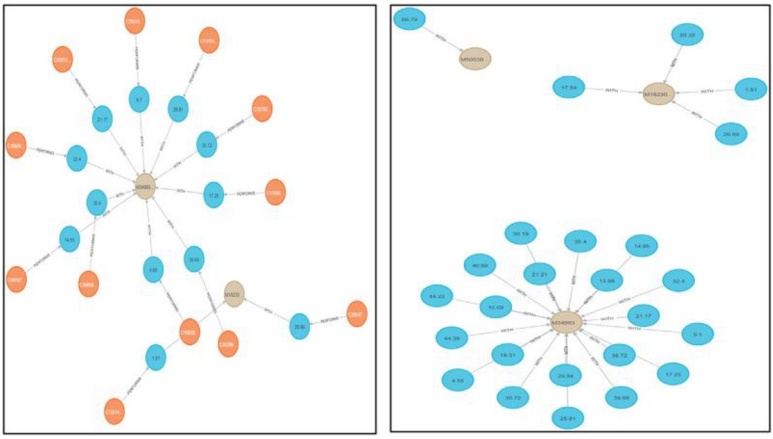

Fig.2. Neo4j Graphs for Relationships in the Dataset

Table 4: Performance of Various Classifiers on the Graph Database

Classifiers	Overall Accuracy (%)		
	Balanced Dataset By EMOTE	**Proposed Model**	**Lifts By**
Random Forest	83.42	**95.38**	**11.96**
Navie Bayes	80.8	**92.87**	**12.07**
XG Boost	90.98	**98.42**	**7.44**
SVM	87.42	**96.67**	**9.25**
KNN	84.8	**94.56**	**9.76**

To exhibit the proposed model`s performance, ROC curve [8] analysis was done with different classifiers such as Navie Bayes, Random Forest, XGBoost, KNN and SVM by utilizing the Bank Sim Graph database. Fig. 3 shows the resultant ROC curves. The ROC curves proves that, classifiers were able to perform well and able to obtain a better accuracy on the graph database that was balanced and constructed by the proposed model. The steep of the curve from the proint 0 to 1 shows and confirms that, the dataset balanced by EMOTE and constructed by the proposed model was the effective dataset in improving the performance of the classifiers.

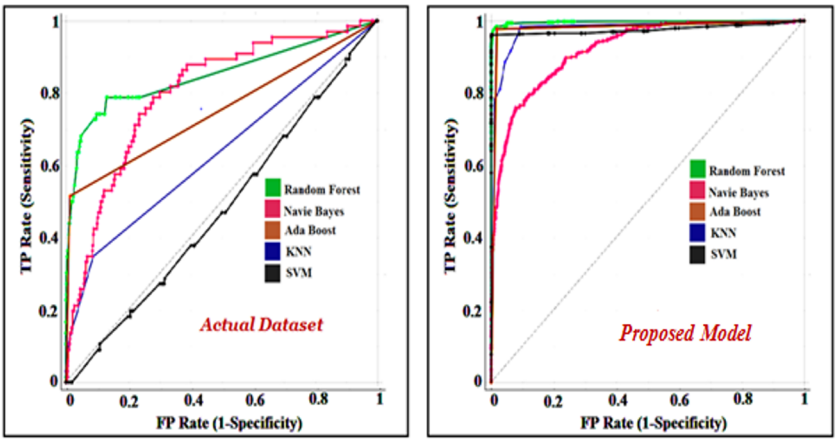

Fig. 3: ROC Comparison of Various Classifiers on Bank Sim Graph Database

As an addition, the AUC [9] [10] analysis is also done. The AUC values obtained from the ROC curves were considered for the analysis. The table 5 shows the AUC values of the different classifiers on the graph database that was refined by the proposed model. These values denote that, the classifiers fall in the range between 0.8 and 0.9 on the actual dataset. This show that, the classifiers predicts maximum number of positive instance as negative instances. Which in turn produces, a huge error on sensitivity and specificity. Whereas the AUC values of the classifiers on the graph database that was refined by the proposed model fall in the range between 0.92 and 0.98. Which consequently reveals that error rate on sensitivity and specificity is very less. This is the important confirmation to prove that, as the relative distribution of the classes in the dataset becomes imbalanced, the model has better influence on the performance of the classifier.

Table 5: AUC Values of the Classifiers on the Graph Database

Classifiers	AUC	
	Balanced Dataset By EMOTE	**Proposed Model**
Random Forest	0.834	**0.953**
Navie Bayes	0.808	**0.928**
XG Boost	0.912	**0.984**
SVM	0.823	**0.966**
KNN	0.854	**0.945**

5 Conclusion

The proposed model effectively addresses the class imbalance issue in the Bank Sim dataset, ensuring a more balanced representation of different classes. This balance allows machine learning classifiers to perform optimally, improving their ability to differentiate between fraudulent and legitimate transactions. To further enhance fraud detection capabilities, a graph-based dataset was constructed using Neo4j, where transactional relationships were represented as a network. This transformation enabled a more structured analysis of fraudulent patterns, capturing complex interactions that traditional tabular datasets might overlook. By applying various machine learning classifiers on the graph dataset, a significant improvement in classification accuracy was observed. The model leveraged graph-based features such as node centrality, community detection, and shortest paths to uncover hidden relationships indicative of fraud. Classifiers trained on this enriched dataset demonstrated enhanced performance in detecting fraudulent activities, reducing false positives and false negatives. Thus, the integration of class balancing techniques, graph database modeling, and advanced machine learning algorithms collectively contributes to a more robust fraud detection system. This approach provides financial institutions with a powerful tool to proactively detect and prevent fraudulent transactions, thereby strengthening security and reducing financial losses.

References

1. Soroor Motie, Bijan Raahemi (2024). Financial fraud detection using graph neural networks: A systematic review, Expert Systems with Applications, Volume 240.
2. Albone, A. (2024). Optimization of fraud detection model with hybrid machine learning and graph database. JURNAL EMACS (Engineering, Mathematics and Computer Science), 6(1), 13–17.
3. Patil, A., Mahajan, S., Menpara, J., Wagle, S., Pareek, P., & Kotecha, K. (2025). Enhancing fraud detection in banking by integration of graph databases with machine learning., MethodsX (Elsevier).
4. Paul, S., Mitra, A., & Koner, C. (2019). A review on graph database and its representation. Proceedings of the International Conference on Recent Advances in Energy-efficient Computing and Communication (ICRAECC), 2019, 1-6.
5. Babu, S., & Ananthanarayanan, N. R. (2017). EMOTE: Enhanced Minority Oversampling TEchnique. Journal of Intelligent & Fuzzy Systems, 33(1), 67-78.
6. B. Can, A.G. Yavuz, E.M. Karsligil, M. Amac Guvensan, A closer look into the characteristics of fraudulent card transactions, IEEE Access 8 (2020) 166095–166109.
7. C. Wang, H. Zhu, Representing fine-grained co-occurrences for behavior-based fraud detection in online payment services, IEEE Trans. Dependable Secur. Comput.19 (1) (2020) 301–315.
8. R. Bin Sulaiman, V. Schetinin, P. Sant, Review of machine learning approach on credit card fraud detection, Hum. Centric Intell. Syst. 2 (1–2) (2022) 55–68.
9. J. Nanduri, Y.W. Liu, K. Yang, Y. Jia, Ecommerce fraud detection through fraud islands and multi-layer machine learning model, in: Proceedings of the 2020 Future of Information and Communication Conference (FICC), Advances in Information and Communication, Springer International Publishing, 2020, pp. 556–570.Volume 2.
10. S. Khan, A. Alourani, B. Mishra, A. Ali, M. Kamal, developing a credit card fraud detection model using machine learning approaches, Int. J. Adv. Comput. Sci.Appl. 13 (2022) 411–418.