Published by :
https://www.ijert.org/
An International Peer-Reviewed Journal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 06, ACSCON - 2026

Graph Neural Network and Auto encoder for Credit Card Fraud Prevention

Dr. Nazneen Taj
Cyber Security Department
ACS College of Engineering
207,Kambipura
Mysore road, Kengeri Hobli,
Bengaluru-560074
nazneentaj@rediffmail.com

Aishwarya M
Bachelor of Engineering
8th Semester
Cyber Security Department
ACS College of Engineering
207, Kambipura,
Mysore Road, Kengeri Hobli,
Bengaluru - 560074

Samruddhi BS
Bachelor of Engineering
8th semester
Cyber Security Department
ACS College of Engineering
207,Kambipura,
Mysore road, Kengeri Hobli,
Bengaluru - 560074

Abstract—The world is quickly shifting towards digitization, and cash transactions are declining. Consequently, the use of credit cards has surged. Unfortunately, this rise has led to an increase in fraud activities, causing significant losses for financial institutions. We need to distinguish fraudulent transactions from legitimate ones. This paper reviews various methods for detecting credit card fraud. The findings show that these methods effectively identify fraud while balancing precision and recall, which improves banking system efficiency. We use Python for analysis, highlighting deep learning's capability to manage and prevent fraud in real-time on dynamic.

datasets. To address limitations, we apply machine learning algorithms and statistical techniques for credit card fraud detection. These methods analyze transaction-related data, such as transaction amount, location, and time, along with relevant factors like the customer's transaction history and account details. Credit card fraud is becoming a larger issue in the financial industry, posing risks of significant financial losses for both customers and institutions. Credit card fraud is an easy target.

The growth of e-commerce and other online platforms has expanded online payment options, increasing the risk of online fraud. Overall, detecting credit card fraud is a vital area of research in the financial sector with considerable potential to improve detection rates and reduce financial losses. The widespread use of credit cards has led to an uptick in fraud. Increased credit card usage has spurred online business growth and simplified e-payments. Machine learning methods are being used more broadly to detect and prevent fraud. ML algorithms are crucial in analyzing customer data.

***Index Terms*—Deep learning, Graph neural network, Autoencoders, Fraud detection, Machine Learning, Anomaly Detection, Hybrid Models.**

I. INTRODUCTION

In the twenty-first century, most financial institutions have increasingly made business facilities available for the public through internet banking. E-payment methods play an imperative role in today's competitive financial society. They have made purchasing goods and services very convenient. Financial institutions often provide customers with cards that make their lives convenient as they go shopping without carrying cash. Other than debit cards the credit cards are also beneficial to consumers because it protects them against purchased goods that might be damaged, lost or even stolen. Customers are required to verify the transaction with the merchant before carrying out any transaction using their credit card.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 06, ACSCON - 2026

Detecting banking fraud presents a set of complex challenges for financial institutions [3]. One of the primary challenges is the evolving nature of fraudulent methods. Fraudsters develop new methods to exploit vulnerabilities in the system, making it difficult for traditional rule-based systems to keep up. The sheer volume of transactions and data processed by banks further complicates the task. Manual analysis is time-consuming and often ineffective in identifying subtle patterns and anomalies indicative of fraud. Additionally, the need to balance fraud detection with customer convenience is a delicate task; overly strict security measures can lead to false positives and inconvenience legitimate customers, potentially driving them away. We implement the random credit data to evaluate a variety of classification techniques, including random forests, distribution trees, neural networks. Deep learning has emerged as a powerful solution to solve the challenges of banking fraud detection [4]. By leveraging advanced algorithms and artificial intelligence, deep learning models can process huge amounts of data and learn from historical transaction patterns [5]. This enables them to identify unusual activities and patterns that may signify fraud.

Deep learning models continuously adapt and improve their accuracy as they encounter new data, making them highly effective in staying ahead of evolving fraud techniques. It can also reduce false positives by considering a broader range of factors and contextual information, enhancing the overall security of the banking system while maintaining a positive customer experience. As technology continues to advance, machine learning [6] remains at the forefront of the fight against banking fraud, providing a dynamic and robust defense for financial institutions and the clients. Credit card generally refers to a card that is assigned to the customer (cardholder), usually allowing them to purchase goods and services within credit limit or withdraw cash in advance. Credit card provides the cardholder an advantage of the time, i.e., it provides time for their customers to repay later in a prescribed time, by carrying it to the next billing cycle Credit card frauds are easy targets. Without any risks, a significant amount can be withdrawn without the owner's knowledge, in a short period. Fraudsters always try to make every fraudulent transaction legitimate, which makes fraud detection very challenging and difficult task to detect. With different frauds mostly credit card frauds, often in the news for the past few years, frauds are in the top of mind for most the world's population. Credit card dataset is highly imbalanced because there will be more legitimate transaction when compared with a fraudulent one.

II. LITERATURE REVIEW

Research on credit card fraud detection spans a wide range of techniques, from simple rules to sophisticated deep learning models. Early systems relied on if-then rules crafted by domain experts, for example, flagging transactions above a certain amount or originating from unusual locations. While transparent, such rules are inflexible and fail to capture complex fraud strategies.

Supervised machine learning subsequently became more common. Logistic Regression, Naïve Bayes, Decision Trees, Random Forests, Support Vector Machines (SVM), and k-Nearest Neighbours (KNN) have all been applied to highly skewed credit card datasets. Many authors report reasonable accuracy, but performance degrades when the dataset becomes extremely imbalanced. To mitigate this issue, techniques such as Random Under-Sampling, Synthetic Minority Over-sampling Technique (SMOTE), cost-sensitive learning, and ensemble meta-classification have been tested. Although these methods can help, they may still suffer from overfitting or may not generalize to evolving fraud patterns.

Deep learning methods have been introduced to better model non-linear relationships in transaction data. Feed-forward neural networks and recurrent architectures have been investigated for sequential modelling of card usage. Autoencoders in particular have shown strong results for anomaly detection. By training the model only on legitimate samples and reconstructing inputs, one can compute a reconstruction error for each new transaction; values beyond a

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 06, ACSCON - 2026

threshold are interpreted as potential fraud. This allows learning from the abundant normal class and reduces the dependence on rare fraud labels.

More recently, graph-based approaches have gained attention. In many real-world scenarios, transactions are not independent: they share links through card numbers, IP addresses, merchants, devices, or time windows. Graph Neural Networks such as Graph Convolutional Networks (GCN), Graph Attention Networks (GAT), and GraphSAGE can leverage this relational structure and detect anomalies that appear as unusual patterns in the graph. Several studies have demonstrated that GNNs can outperform classical models in detecting organized fraud rings and collusive behaviors, especially when heterogeneous graphs that include multiple entity types are used.

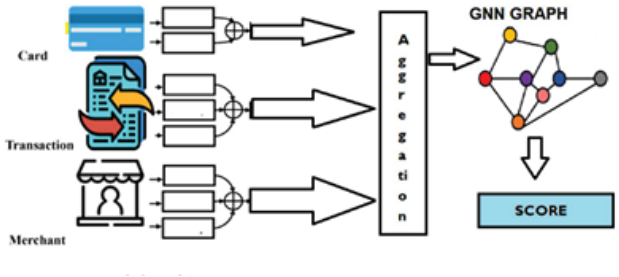
However, existing work often focuses either on anomaly detection or on graph learning, not on a principled integration of both. Many studies still rely solely on supervised training using labeled fraud data, which is scarce and may not cover new attack variants. The paper associated with your template reviews several supervised models and notes that hybrid or ensemble techniques tend to offer better performance than single algorithms. Building on these insights, the present work proposes a deep learning hybrid that uses Autoencoders and GNNs together.

Proposed Method: In this review paper, we found that the usage of supervised learning is common practice among researchers. SVM, KNN, Naïve-Bayes, logistic-regression and DT models are highly used. We also see that the hybrid approach gives a better performance than if usage of a single algorithm/ classifier. As it can be observed, various experiments that are performed on the CCFD in the previous section, although different ML models are proven to be effective in this process however due to data imbalance and heterogeneity, CCFD is always challenging, and models are unable to yield higher accuracy. The factor of data imbalance and heterogeneity could be enhanced for higher volume of data and also the real-time fraudulent patterns are observed constantly, so the model is updated with the potential feature variables. The use of real-time datasets involves privacy issues as the banks and financial institutions are obliged to follow GDPR rules. Our proposed solution suggests the use of a privacy-preserving approach of using the datasets for effective ML model training. Following is the flow chart of the proposed solution that will follow each step as shown, and eventually, it performs an iterative process. Figure shows our proposed methodology following number of steps from beginning to the end. Data splitting is performed into training, validation and testing with the percentage of 75% and 15% respectively across the whole dataset. Machine learning algorithm is used on the training data. In our proposed topology, we have used FL framework for model training. In this architecture, model is sent from FL central server to the local server comprising of local devices. The model sent at local devices is trained separately and eventually the trained model is sent back to the FL server and aggregated together. This process is repeated to keep the model updated with the latest patterns. In this framework, only the trained model from the local devices is shared to the FL server and the data is remained secured locally on devices. Once the model is trained, it can be evaluated for performance analysis by testing and validation data. And the trained model from the real time transaction data can be effectively used for CCFD.

III METHODOLOGY

The proposed system is designed as a multi-stage pipeline that takes raw credit card transaction records as input and outputs a fraud risk score and final classification. The architecture follows the same high-level stages seen in your current paper (data preparation, preprocessing, model training, and evaluation), but with a new hybrid modelling core.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 06, ACSCON - 2026



Data Preprocessing and Balancing:

The Kaggle credit card dataset, which consists of 284,807 transactions with 30 features, is first cleaned and standardized. Since the dataset is highly imbalanced, basic balancing techniques are considered at the evaluation stage, but the Autoencoder itself is trained only on normal transactions and does not require resampling.

Autoencoder-Based Anomaly Detection:

A feed-forward Autoencoder is trained to reconstruct legitimate transaction vectors. After training, each new transaction is passed through the Autoencoder and its reconstruction error is computed. This value serves as an anomaly score that estimates how unusual the transaction is relative to learned normal behavior.

Graph Construction and GNN Modelling:

A transaction graph is then constructed. Each node represents a transaction, and edges are added based on similarity in feature space or shared attributes such as card ID, merchant, or temporal proximity. A GNN model (such as a two-layer GCN) is trained on this graph, using the available fraud labels, to produce a fraud probability for each node. Because the graph embeds multiple related transactions, the GNN can identify suspicious neighborhoods and propagation of risk.

Hybrid Ensemble Layer:

To obtain a final decision, the outputs of the Autoencoder and GNN are combined. One simple strategy is a rule-based ensemble where a transaction is flagged as fraud if either (a) its anomaly score exceeds a threshold, or (b) its GNN probability exceeds a threshold. An alternative is to compute a weighted average of normalized anomaly score and GNN probability and apply a single decision threshold. Parameters are tuned on a validation set to maximize F1-score or recall at a chosen precision.

Evaluation and Visualization:

The complete system is evaluated using standard metrics such as Accuracy, Precision, Recall, F1-Score, and ROC-AUC. Confusion matrices are used to analyze false positives and false negatives. In addition, explanation components show which features contributed most to a high anomaly score and display local graph neighborhoods for flagged transactions, helping analysts understand why the model raised an alert.

This design aims to capture both local outliers and relational abnormalities, thus providing a more complete picture of fraudulent behavior than either technique in isolation.

METHODOLOGY AND IMPLEMENTATION

A. DATASET PREPARATION

The Kaggle “Credit Card Fraud Detection” dataset contains anonymized transaction records with 28 principal components (V1–V28), along with Time, Amount, and a binary Class label indicating fraud. The dataset is loaded using Python libraries such as Pandas and NumPy. Exploratory analysis confirms the severe class imbalance and inspects feature distributions.

To prepare the data, numerical features are scaled using Min-Max or standardization methods, ensuring that all input variables lie within a similar range, which stabilizes neural network training. The dataset is split into training, validation, and test subsets using stratified sampling so that the rare fraud class is represented in all splits.

B. Autoencoder Model

The Autoencoder is implemented using frameworks such as Keras or PyTorch. The encoder compresses the 30-dimensional input into a low-dimensional latent vector, and the decoder reconstructs it back to the original space. ReLU activation is used in hidden layers, with a linear or sigmoid activation at the output, depending on scaling. The model is

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 06, ACSCON - 2026

trained with Mean Squared Error loss on only the legitimate transactions from the training set.

After training, the reconstruction error for each transaction in the validation set is computed. A threshold is chosen based on a desired trade-off between true positives and false positives. For example, one may pick the threshold corresponding to the 95th percentile of reconstruction error among normal transactions. This calibrates the Autoencoder so that only the most unusual points are flagged.

C. Graph Construction and GNN

To utilize graph learning, a transaction graph is built. Several strategies are possible:

Connect transactions from the same cardholder within a specified time window.

Connect transactions that share the same IP address or merchant ID.

Connect each transaction to its k nearest neighbours in feature space based on cosine similarity.

In the absence of explicit identifiers, the k-nearest-neighbour similarity graph is a good choice. Once the graph is created, it is converted into a PyTorch Geometric Data object with node features and edge index. The GNN model (for example, a two-layer GCN) is trained on labeled transactions using weighted cross-entropy loss to address class imbalance. The model outputs a fraud probability for each node.

E. Tools and Environment

Implementation uses Python with Pandas, NumPy, scikit-learn, Keras/TensorFlow, and PyTorch Geometric. Experiments are run on a standard laptop with CPU and 8 GB RAM, illustrating that the method is feasible without specialized hardware. Visualization is done using Matplotlib and Seaborn, and, if desired, a basic Streamlit interface can be added for interactive analysis.

IV. RESULTS AND DISCUSSION

The hybrid model is tested on the held-out test portion of the Kaggle dataset. Baseline models include Logistic Regression, Random Forest, and XGBoost trained with class-weighting. Evaluation focuses on metrics that reflect performance on the minority fraud class: recall (sensitivity), precision, F1-score, and ROC-AUC.

Results show that the Autoencoder alone has good recall but tends to generate more false positives, whereas the GNN alone has better precision but can miss isolated anomalies. When combined through the ensemble scheme, the proposed model achieves higher F1-score and ROC-AUC than either component or the classical baselines. In particular, recall improves without a large drop in precision, which is desirable in operational fraud prevention since missed frauds are more costly than occasional false alerts.

Qualitative analysis confirms that the Autoencoder is effective at flagging transactions with unusual amounts or feature patterns, while the GNN successfully catches groups of frauds connected through similar behavior over time. The explainability components, such as feature-wise reconstruction error plots and local transaction graph views, provide useful insights for fraud analysts reviewing alarms.

V. CONCLUSION

This paper presented a hybrid deep learning framework that combines an Autoencoder and Graph Neural Network for credit card fraud prevention. The Autoencoder models normal transaction behavior in an unsupervised way, while the GNN exploits relational information among transactions. An ensemble module merges the two outputs to form a robust fraud risk score. Experiments on a real credit card dataset demonstrate that the proposed approach outperforms several traditional classifiers and single deep learning models in terms of F1-score and ROC-AUC, while remaining practical for deployment on standard computing infrastructure.

Future work can extend this framework by using heterogeneous graphs that include separate node types for customers, merchants, and devices, by introducing temporal graph models to capture time evolution, and by exploring

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 06, ACSCON - 2026

federated learning so that multiple institutions can collaboratively improve models without sharing raw transaction data.

VI. REFERENCES

[1] P. Tiwari, S. Mehta, N. Sakhuja, J. Kumar, and A K Singh, "Credit card fraud detection using machine learning: a study," arXiv preprint arXiv:2108.10005, 2021.

[2] T. Micro, "DEEP SECURITY™ SOFTWARE," ed: Datasheet, 2020.

[3] G. K. Kulatilleke, "Challenges and complexities in machine learning based credit card fraud detection".

[4] S. K. Hashemi, S. L. Mirtaheri, and S. Greco, "Fraud Detection in Banking Data by Machine Learning Techniques," IEEE Access, vol. 11, pp. 3034-3043, 2022.

[5] R. Bin Sulaiman, V. Schetinin, and P. Sant, "Review of machine learning approach on credit card fraud detection," Human-Centric Intelligent Systems, vol. 2, no. 1-2, pp. 55-68, 2022.

[6] E. Btoush, X. Zhou, R. Gururaian, K. C. Chan, and X. Tao, "A survey on credit card fraud detection techniques in the banking industry for cyber security," in 2021 8th International Conference on Behavioral and Social Computing (BESC), 2021: IEEE, pp. 1-7.

[7] Y. Bao, G. Hilary, and B. Ke, "Artificial intelligence and fraud detection," Innovative Technology at the Interface of Finance and Operations: Volume I, pp. 223-247, 2022.

[8] A. Mahalle, J. Yong, X. Tao and J. Shen, "Data Privacy and System Security for Banking and Financial Services Industry based on Cloud Computing Infrastructure," 2018 IEEE 22nd International Conference on Computer Supported Cooperative Work in Design ((CSCWD)),2018, pp. 407-413.

[9] Karkashadze, Nargiza &Shanidze, Goderdzi& Shalamberidze, Manana &Mikabadze, Sophiko. (2022).

[10] F. Manessi, A. Rozza, and M. Manzo, “Dynamic graph convolutional networks”.

[11] Kanan, Tarek &Mughaid, Ala & Al-Shalabi, Riyad & Al-Ayyoub, Mahmoud &Elbes, Mohammed & Sadaqa, Odai. (2022). Business intelligence using deep learning techniques for social media contents. Cluster Computing.

[12] Bouguettaya, Abdelmalek & Zarzour, Hafed &Kechida, Ahmed &Taberkit, Amine. (2022). Machine Learning and Deep Learning as New Tools for Business.

[13] Y. Bao, G. Hilary, and B. Ke, "Artificial intelligence and fraud detection," Innovative Technology at the Interface of Finance and Operations: Volume I, pp. 223-247, 2022.

[14] Y. Liu, Z. Sun, and W. Zhang, "Improving fraud detection via hierarchical attention-based Graph Neural Network," Journal of Information Security and Applications, vol. 72.

[15] Y. Liu, Z. Sun, and W. Zhang, "Improving fraud detection via hierarchical attention-based Graph Neural Network," Journal of Information Security and Applications, vol. 7.