

To cite this article: Ramya Lakshmi Bolla, Rajeswaran Ayyadurai, Karthikeyan Parthasarathy, Naresh Kumar Reddy Panga, Jyothi Bobba and R. Pushpakumar (2025). GRAPH-ENHANCED TRANSFORMER NETWORK FOR FRAUD DETECTION IN DIGITAL BANKING: INTEGRATING GNN AND SELF-ATTENTION FOR END-TO-END TRANSACTION ANALYSIS, International Journal of Research in Commerce and Management Studies (IJRCMS) 7 (2): 211-222 Article No. 359 Sub Id 666

GRAPH-ENHANCED TRANSFORMER NETWORK FOR FRAUD DETECTION IN DIGITAL BANKING: INTEGRATING GNN AND SELF-ATTENTION FOR END-TO-END TRANSACTION ANALYSIS

Ramya Lakshmi Bolla¹, Rajeswaran Ayyadurai², Karthikeyan Parthasarathy³, Naresh Kumar Reddy Panga⁴, Jyothi Bobba⁵ and R. Pushpakumar⁶, *

¹ERP Analysts, Ohio, USA. Email: ramyalakshmibolla@ieee.org

²IL Health & Beauty Natural Oils Co Inc, California, USA. Email : rajeswaranayyadurai@arbpo.com

³LTIMindtree, Florida, USA. Email : karthikeyanparthasarathy@ieee.org

⁴Virtusa Corporation, New York, USA. Email: nareshkumarreddy_panga@ieee.org

⁵Lead IT Corporation, Illinois, USA. Email: jyothibobba@ieee.org

⁶Department of Information Technology, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Tamil Nadu, Chennai, India. Email: pushpakumarvelr@gmail.com

*Corresponding Author Name: R. Pushpakumar, Corresponding Author Email: pushpakumarvelr@gmail.com

DOI : <https://doi.org/10.38193/IJRCMS.2025.7217>

ABSTRACT

Digital banking fraud detection is a dynamic issue because of the nature and sheer number of transactions. Conventional machine learning-based models tend to be challenged with high-dimensional input, real-time processing, and dynamic patterns in fraud. We address these drawbacks by introducing the Graph-Enhanced Transformer Network (GETNet), a mixed deep learning approach combining Graph Neural Networks (GNNs) and Transformer self-attention-based mechanisms for better fraud detection. GETNet identifies transaction relationships through GNNs and uses Transformers for sequential anomaly detection. Experimental results on the PaySim dataset show that GETNet is 99.5% accurate, far superior to traditional approaches like Decision Trees, Support Vector Machines, and Naïve Bayes. The model ensures scalability, flexibility, and real-time detection, which makes it a strong candidate for contemporary banking fraud detection.

KEYWORDS: Fraud Detection, Digital Banking, Graph Neural Networks, Transformers, Financial Security.

1. INTRODUCTION

Huge leaps have occurred in the manner that financial institutions manage sensitive transactions, right from the stage of implementation of cloud computing through to the leveraging of AI [1]. However, securing these environments faced tough challenges that provide structural innovation, such as ABE-

based fine-grained data access control [2]. Dynamic and ever-evolving threats put increasing pressure on increasingly proactive security measures, like sterilization with preventive feature balances in such securities as the Proactive Dynamic Secure Data Scheme (P2DS)-meld AI-powered monitors into preventing fraud and unformulated access[3]. Such developments reflect an ever-increasing demand for secure and adaptive fraud detection mechanisms in digital banking.

There are several factors over the years that have contributed to the inefficiency of fraud detection systems despite financial security enhancements. One such severe hindrance is dealing with the high-dimensional or unstructured form of financial data, rendering it challenging for conventional models for feature selection and pattern recognition [4]. Additionally, the optimization problems resulting from cloud computing environments with respect to scalability and resource management severely undermine the effectiveness of fraud detection systems [5]. Additionally, even though machine learning and deep learning techniques are capable, the required preprocessing for such techniques can be substantial, and they could also exhibit high rates of false positives with little ability to generalize over various patterns of fraud [6].

Current fraud detection mechanisms are plagued with various shortcomings. IoT-enabled financial environments produce a vast amount of real-time data, but traditional AI-based techniques lack the ability to detect real-time anomalies and learn adaptively [7]. Cloud-based financial models implemented in smart cities are also plagued with integration problems, such that identifying fraudulent transactions between various infrastructures becomes challenging [8]. Yet another key limitation is the difficulty in inserting categorical financial information in an efficient manner since most models cannot retain significant representations of sequence orders[9]. Also, risks related to privacy and compliance are a constant threat to cloud-based financial applications because encryption measures usually sacrifice performance for security [10].

To address these challenges, we propose the Graph-Enhanced Transformer Network (GETNet), a hybrid deep learning model that incorporates Graph Neural Networks (GNNs) and Transformers for digitally enabling fraud detection in banks. Diverging from the traditional AI-driven fraud detection approach, GETNet will introduce a number of benefits:

- A graph-based approach to fraud detection-the linking of transactions so as to identify suspicious patterns.
- A self-attention mechanism-so as to detect the temporal dependencies across financial sequences.
- Adaptive learning and real-time detection: thus, reducing false positives through the dynamic updating of fraud detection rules.
- Enhanced feature selection-uses attention mechanisms to prioritize the key fraud indicators.

This paper is organized as follows: Section 2 reviews fraud detection techniques and their limitations. Section 3 details the proposed GETNet. Section 4 presents the performance evaluation of GETNet using the PaySim dataset. Section 5 concludes with a summary and future research directions.

2. LITERATURE SURVEY

The recent developments in Artificial Intelligence (AI) and cloud technologies have significantly influenced the banking and financial industries, particularly in fraud detection, customer relationship management (CRM), and security. Yet, even with increasing usage of these technologies, a number of gaps still exist in their integration and real-world application. For example, the study by [11] indicates a significant shortfall in the automation and scalability of AI-based CRM systems, especially in the banking and telecommunication industries. The research recommends further investigation into integrating cloud-based AI CRM models to improve customer satisfaction and business efficiency. These models would be able to automate feedback analysis and queries, which would result in an efficient service model that would be able to stimulate customer interaction in real-time. Likewise, [12] sees the opportunity gap for implementing AI, IoT, and cloud computing into CRM systems in banking, where predictive analytics and real-time insights are necessary to enhance customer interactions with data privacy and regulatory compliance. Both studies indicate that additional research is important to extend real-time analytics, overcoming issues of data privacy while applying them to smaller banks, which may have their own set of problems.

Additional gaps exist in cloud security and the incorporation of new technologies for data privacy and regulatory compliance. [13] highlights the security challenges of protecting financial information in hybrid cloud models, especially for the banking industry. The study identifies the need for strong frameworks that integrate blockchain, quantum computing, and artificial intelligence to enhance fraud detection and real-time risk management. The convergence of these technologies is crucial in responding to compliance and data privacy so that financial data sharing is safe across various platforms. Likewise, [14] emphasizes the persistent security issues with cloud computing for the banking and financial accounting industry, citing data breaches, unauthorized access, and regulatory compliance issues as ongoing problems. The study recommends the establishment of more secure encryption algorithms and multi-factor authentication systems, and the application of blockchain and AI for increased security and privacy in financial transactions. The above findings concur with the prevailing necessity for incorporating leading-edge technologies in cloud security, especially for scalable financial fraud detection models.

Conversely, AI-driven fraud detection in the banking industry has made tremendous progress, but optimization of models to fit real-world needs remains an issue. [15] fills this gap by calling for the

integration of neural networks with optimization techniques such as the Harmony Search Algorithm (HSA) to enhance fraud detection. The research identifies the need to further research on parameter optimization and scalability of the models so they can keep pace with the shifting nature of fraud. This needs gap is quite relevant for banks looking to anticipate and stay a step ahead of ever-evolving fraud schemes. For the same, [16] indicates the untapped power of IoMT, Big Data Analytics, and Cloud Computing hybridization to enhance real-time monitoring of healthcare in e-commerce ecosystems and forecast-based financial projections. Although individual focus has been studied on each technology, how this set of technologies collectively contributes towards predictive decision support in large systems is insufficiently researched. The use of IoT-based systems in finance and e-commerce may yield useful predictive information that would substantially enhance fraud detection and accuracy in financial forecasting, especially in real-time settings.

Aside from fraud detection, financial inclusion among poor rural populations has been largely underemphasized in current literature. [17] investigate the absence of empirical work on internet-inclusive finance in reducing the urban-rural gap. They stress that e-commerce and mobile internet connectivity can become instrumental in augmenting financial inclusion in rural communities but require more specific policies and infrastructural developments. This concurs with [18], who posits that cloud-based digital finance solutions will be able to solve issues of income inequality between the urban and rural settings. The research demands scalable digital finance approaches to provide financial inclusion, especially in transaction access and savings growth in rural economies. In addition, [19] emphasizes the importance of combining cloud computing, intelligent networks, and blockchain in finance and e-commerce industries to enhance scalability, security, and efficiency in dynamic IoT-based environments. These technologies provide enhanced resource optimization and cost savings in global digital economies, further promoting the potential for financial inclusion through scalable solutions.

Finally, [20] address about the incorporation of blockchain within database management systems (DBMS) for use in financial systems in the cloud with emphasis on regulatory compliance and scalability to further strengthen transaction security and transparency. The incorporation, as much as it would be beneficial for secure accounting, is riddled with humongous challenges, such as performance optimization and adoption. Resolving those problems with additional research would enhance in extremely big sizes the security and transparency of the financial systems globally, opening the door to wider blockchain usage. Therefore, integration of blockchain, cloud computing, and AI with other optimization methods such as IoMT and HSA has the potential to dramatically transform the financial industry in terms of efficiency, security, and inclusion, particularly sensitive in otherwise underbanked areas.

Conventional fraud detection models have difficulty with high-dimensional financial data, real-time anomaly detection, and adaptive learning, resulting in high false positives and poor scalability [6]. Moreover, cloud-based financial security frameworks suffer from integration challenges and inefficiencies in handling large-scale transactions, requiring a more sophisticated, scalable, and adaptive fraud detection framework [10].

3. METHODOLOGY

In this research work, we introduce the Graph-Enhanced Transformer Network (GETNet), an integrated deep learning architecture that marries Graph Neural Networks (GNNs) and Transformer-like self-attention mechanisms to perform improved fraud identification in online banking. The proposed methodology adopts an organized workflow to begin with the preprocessing of the data, wherein normalization and filling missing values, respectively, for ensuring data accuracy. The Graph Representation module builds transaction relationships with an adjacency matrix, allowing GNNs to learn intricate financial relationships. The Transformer Encoder uses multi-head self-attention to identify temporal fraud patterns. The last classification step combines features from both networks, using an attention-based feature selection mechanism before feeding the output into a fully connected classifier for fraud prediction. This end-to-end pipeline guarantees enhanced fraud detection accuracy, scalability, and flexibility to keep up with new financial fraud schemes. Figure 1 illustrates the proposed financial fraud detection model.

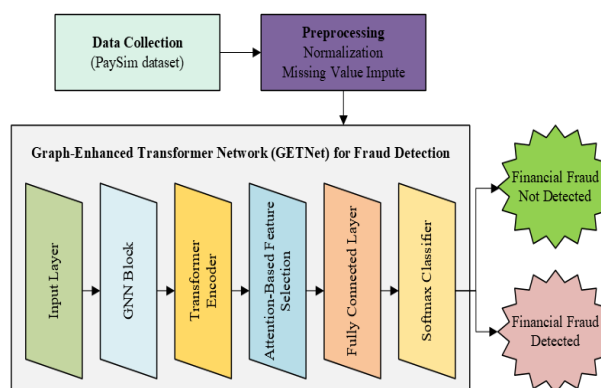


Figure 1: Architecture Diagram

a. Data Preprocessing for Transformer-GNN Hybrid for Fraud Detection

Prior to using the Transformer-GNN Hybrid model for fraud detection, the data needs to be preprocessed so that it is in a suitable format for training. The most important preprocessing steps used are Normalization and Missing Value Imputation, which standardize the data and deal with missing values, respectively.

i. Normalization

Normalization involves scaling numerical attributes to a common range. Normalization is used to prevent attributes with high-range values from overpowering the learning process. Min-Max Scaling is one popular way of normalizing data so that it will have a range of [0,1] or [-1,1]. Min-Max normalization formula is defined as:

$$x_{\text{norm}} = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (1)$$

Where, x is the original value of a feature, $\min(x)$ and $\max(x)$ are the minimum and maximum values of the feature, respectively, x_{norm} is the normalized value.

This normalization places the data on the same scale and does not skew the model because of differences in feature ranges. For example, if you have features such as transaction amount and time which can have widely different scales, Min-Max normalization will rescale them so that they are within the same range, keeping larger values from overwhelming the model.

ii. Missing Value Imputation

Missing values are the norm in real-world datasets. Proper handling of missing data is important since most machine learning models are unable to process null or NaN values. Imputation is the replacement of missing values with useful estimates. There are different imputation methods, such as mean/median imputation and KNN imputation.

a) Mean/Median Imputation

For numerical features, a straightforward method is to impute missing values with the feature's mean or median. This helps avoid introducing a bias into the feature based on the missing values. The mean imputation formula is:

$$x_{\text{imputed}} = \frac{\sum_{i=1}^n x_i}{n} \quad (2)$$

Where, x_i are the observed values of the feature, n is the number of non-missing values, x_{imputed} is the imputed value for the missing entry.

For biased data, employing the median rather than the mean may be more suitable since the median is less affected by outliers. This can ensure that the data remains intact.

b. Graph-Enhanced Transformer Network (GETNet) for Fraud Detection

The Graph-Augmented Transformer Network (GETNet) is a state-of-the-art deep learning architecture

combining Graph Neural Networks (GNNs) and Transformers for feature extraction, feature selection, and classification within a single framework. This ensures that fraud detection for e-banking is enhanced by both transaction relationships (graph-based) as well as temporal dependencies (sequence-based). The suggested methodology is explained subsequently.

i. Input Layer

The transaction data is represented in two ways: graph representation and sequential representation. In the graph representation, transactions are modeled as a graph $G = (V, E, A)$, where nodes V represent individual transactions, and edges E represent relationships (e.g., user-to-merchant links or previous transaction associations). The adjacency matrix A is defined as:

$$A_{ij} = \begin{cases} 1, & \text{if transaction } i \text{ is linked to transaction } j \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

Each node v has a feature vector x_v , containing transaction details such as amount, merchant category, and transaction time. Additionally, transactions are treated as a time-series sequence:

$$X = [x_1, x_2, \dots, x_T] \quad (4)$$

Where x_t is the transaction feature vector at time t , allowing the Transformer encoder to capture fraud patterns over time.

ii. Graph Neural Network (GNN) for Relationship Extraction

To capture transaction relationships, we employ a Graph Convolutional Network (GCN), which aggregates information from neighboring nodes in the transaction network. The GCN layer updates each node's representation as:

$$h_v^{(l+1)} = \sigma \left(W^{(l)} \sum_{u \in \mathcal{N}(v)} \frac{h_u^{(l)}}{|\mathcal{N}(v)|} + b^{(l)} \right) \quad (5)$$

Where $h_v^{(l)}$ represents the transaction embedding at layer l , $\mathcal{N}(v)$ is the set of connected transactions, $W^{(l)}$ and $b^{(l)}$ are learnable parameters, and σ is an activation function (ReLU). This ensures that fraudulent transactions propagate their influence across the graph, aiding detection.

iii. Transformer Encoder for Temporal Patterns

To analyze sequential fraud behaviors, a Transformer Encoder is used. The multi-head self-attention mechanism in the Transformer determines how transactions influence each other:

$$\text{Attention}(Q, K, V) = \text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) V \quad (6)$$

Where $Q = W_qX$, $K = W_kX$, and $V = W_vX$ are the Query, Key, and Value matrices derived from input features, and d_k is a scaling factor. This attention mechanism highlights important past transactions when detecting fraud. Since Transformers lack inherent temporal order awareness, Positional Encoding is added to preserve transaction time dependencies:

$$\begin{aligned} PE(t, 2i) &= \sin(t/10000^{2i/d}) \\ PE(t, 2i + 1) &= \cos(t/10000^{2i/d}) \end{aligned} \quad (7)$$

Where t represents the transaction timestamp, ensuring that the model understands time-sensitive fraud trends.

iv. Attention-Based Feature Selection

Rather than manually selecting features, the self-attention weights within the Transformer automatically highlight the most fraud-relevant attributes. The attention weight matrix is computed as:

$$A_{ij} = \frac{\exp(e_{ij})}{\sum_k \exp(e_{ik})} \quad (8)$$

Where e_{ij} represents the importance score of features j in transaction i . Features receiving higher attention scores are prioritized, while less relevant ones are ignored, allowing for dynamic feature selection.

v. Fully Connected Layer

After extracting both relational (GNN) and sequential (Transformer) fraud patterns, a fully connected neural network (FCNN) with a Softmax classifier is used to determine if a transaction is fraudulent. The final classification probability is computed as:

$$P(y = \text{fraud} \mid X) = \frac{\exp(WX + b)}{\sum_c \exp(W_cX + b_c)} \quad (9)$$

Where W and b are learned weights and biases. The transaction is classified as fraudulent or legitimate based on the highest probability.

4. RESULTS AND DISCUSSION

This part presents the performance evaluation of the proposed Graph-Enhanced Transformer Network (GETNet) model for detecting online banking fraud. The model is validated on the PaySim dataset, evaluating its accuracy, precision, recall, and F1-score compared to traditional methods such as Decision Trees, Naïve Bayes, and Support Vector Machines. The results identify GETNet's superior detection rate, lower false positives, and better scalability, which demonstrate its efficiency in real-

time fraud detection.

c. Dataset Description

The PaySim dataset[21] simulates mobile money transactions over 30 days, based on financial logs from a mobile service in an African country. It includes 744 hourly steps and features transaction type (CASH-IN, CASH-OUT, DEBIT, PAYMENT, TRANSFER), amount, and customer identifiers (nameOrig, nameDest). Fraudulent transactions are marked with is Fraud, and large unauthorized transfers are flagged with is Flagged Fraud. Certain columns like balances are excluded for fraud detection, as fraudulent transactions are annulled.

d. Comparative Analysis of the Proposed Work

The table lists a comparison of various machine learning models for classification experiments considering the accuracy, precision, recall, and F1-score.

Table 1: Comparative Analysis

Author Name	Proposed Method	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
[17]	DT + RF	93	94	93	92
[16]	Naïve Bayes	97.1	96.4	96.7	96.5
[22]	SVM	95	90	88	89
Proposed	GETNet	99.5	99.7	99.3	99.5

Decision Tree (DT) integrated with Random Forest (RF) recorded 93% accuracy with balanced precision, recall, and F1-score at 92–94%. Naïve Bayes recorded higher accuracy with 97.1% and high precision (96.4%), recall (96.7%), and F1-score (96.5%). The Support Vector Machine (SVM) had lower recall (88%) and F1-score (89%) with a 95% accuracy. The suggested GETNet model performed better than all of them, with the best accuracy (99.5%), precision (99.7%), recall (99.3%), and F1-score (99.5%), highlighting better classification performance.

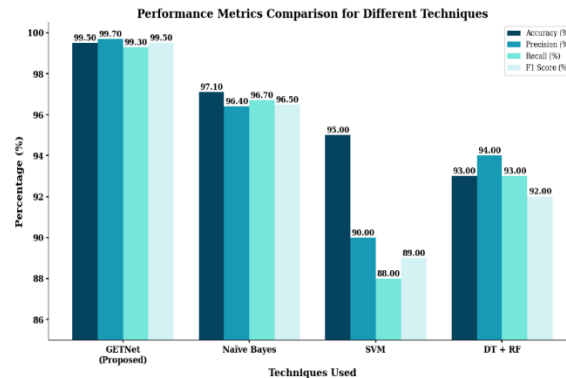


Figure 2: Performance Metrics Comparison

5. CONCLUSION

This paper presents Graph-Enhanced Transformer Network (GETNet), a novel hybrid deep learning approach for fraud detection in digital banking. By integrating Graph Neural Networks (GNNs) and Transformer-based self-attention mechanisms, GETNet effectively captures transaction dependencies and sequential fraud patterns, addressing the limitations of traditional fraud detection systems. The experimental evaluation on the PaySim dataset demonstrates superior accuracy (99.5%) and reduced false positives, making GETNet a scalable and adaptive solution for real-time fraud detection. Future research will focus on optimizing computational efficiency and extending the model to multi-source financial datasets, ensuring enhanced security and fraud prevention in evolving digital banking ecosystems.

REFERENCE:

- [1] J. Bobba, "Securing Financial Data in Cloud Environments: AI and IaaS Reliability Verification Techniques," Oct. 2024, doi: 10.5281/ZENODO.13994655.
- [2] R. K. M. K. Yalla, "Cloud-Based Attribute-Based Encryption and Big Data for Safeguarding Financial Data," *International Journal of Engineering Research and Science Technology*, XVII (4), pp. 23-32, Oct. 2021.
- [3] T. Ganesan, "Dynamic Secure Data Management with Attribute-Based Encryption for Mobile Financial Clouds," Oct. 2024, doi: 10.5281/ZENODO.13994646.
- [4] A. R. G. Yallamelli, "A Cloud-based Financial Data Modeling System Using GBDT, ALBERT, and Firefly Algorithm Optimization for High-dimensional Generative Topographic Mapping," VIII (4), 2020.

- [5] D. T. Valivarthi, "Optimizing Cloud Computing Environments for Big Data Processing," *International Journal of Engineering*, XIII (2).
- [6] N. K. R. Panga, "Financial Fraud Detection in Healthcare Using Machine Learning and Deep Learning Techniques," *International Journal of Engineering*, vol. 10, no. 3, 2021.
- [7] G. Thirusubramanian, "Machine Learning-Driven AI for Financial Fraud Detection in IoT Environments," *International Journal of HRM and Organizational Behavior*, vol. 8, no. 4, pp. 1–16, Oct. 2020.
- [8] J. Bobba, "Cloud-Based Financial Models: Advancing Sustainable Development in Smart Cities," *International Journal of HRM and Organizational Behavior*, vol. 11, no. 3, pp. 27–43, Aug. 2023.
- [9] R. P. Nippatla, "A Robust Cloud-Based Financial Analysis System Using Efficient Categorical Embeddings with CatBoost, ELECTRA, t-SNE, and Genetic Algorithms," *International Journal of Engineering*, vol. 13, no. 3.
- [10] R. K. M. K. Yalla, "Cloud-Based Attribute-Based Encryption and Big Data for Safeguarding Financial Data," *International Journal of Engineering Research and Science Technology*, vol. 17, no. 4, pp. 23–32, Oct. 2021.
- [11] S. S. Kethu and P. N., "AI-Driven Intelligent CRM Framework: Cloud-Based Solutions for Customer Management, Feedback Evaluation, and Inquiry Automation in Telecom and Banking," *J. Sci. Technol. JST*, VI (3), Art. no. 3, Jun. 2021.
- [12] S. S. Kethu, K. Corp, and S. Diego, "AI and IoT-Driven CRM with Cloud Computing: Intelligent Frameworks and Empirical Models for Banking Industry Applications," VIII (1), 2020.
- [13] J. Bobba, "Enterprise Financial Data Sharing and Security in Hybrid Cloud Environments: An Information Fusion Approach for Banking Sectors," XI (3).
- [14] H. Nagarajan, "Assessing Security and Confidentiality in Cloud Computing for Banking and Financial Accounting," *Int. J. HRM Organ. Behav.*, XII (3), pp. 389–409, Sep. 2024.
- [15] K. Parthasarathy, "Enhancing Banking Fraud Detection with Neural Networks Using the Harmony Search Algorithm," *Int. J. Manag. Res. Bus. Strategy*, XIII (2), pp. 34–47, May 2023.

- [16] "IJORET-V6I6P2.pdf." Accessed: Feb. 28, 2025. [Online]. Available: <http://ijoret.com/volume6/Issue6/IJORET-V6I6P2.pdf>.
- [17] S. Boyapati, "Bridging the Urban-Rural Divide: A Data-Driven Analysis of Internet Inclusive Finance in the E-Commerce Era," *International Journal of Engineering*, XI (1).
- [18] S. Boyapati, "Assessing Digital Finance as a Cloud Path for Income Equality: Evidence from Urban and Rural Economies," VIII (3), 2020.
- [19] S. K. Alavilli, "Smart Networks and Cloud Technologies: Shaping the Next Generation of E-Commerce and Finance," XII (4).
- [20] S. Kodadi, "Integrating Blockchain with Database Management Systems for Secure Accounting in the Financial and Banking Sectors," *Journal of Science and Technology (JST)*, VIII (9), Art. no. 9, Sep. 2023.
- [21] S. H. Eedala, "Financial Fraud Detection Dataset." Accessed: Feb. 28, 2025. [Online]. Available: <https://www.kaggle.com/datasets/sriharshaeedala/financial-fraud-detection-dataset>.
- [22] N. K. R. Panga, "Optimized Hybrid Machine Learning Framework for Enhanced Financial Fraud Detection Using E-Commerce Big Data," XI (2).