Indian Journal of Computer Science and Technology
https://www.doi.org/10.59256/indjcst.20250403008
Volume 4, Issue3 (September-December 2025), PP: 39-44.
www.indjcst.com
ISSN No: 2583-5300

Enhancing Financial Fraud Detection by Leveraging Llama2 NLP and Neo4j

Shaik Mohammad Adil[1], Dr. Mohd Rafi Ahmed[2]
[1]*Student, MCA Deccan College of Engineering and Technology, Hyderabad, Telangana, India.*
[2]*Associate Professor, MCA Deccan College of Engineering and Technology, Hyderabad, Telangana, India.*

To Cite this Article: *Shaik Mohammad Adil[1], Dr. Mohd Rafi Ahmed[2], "Enhancing Financial Fraud Detection by Leveraging Llama2 NLP and Neo4j", Indian Journal of Computer Science and Technology, Volume 04, Issue 03 (September-December 2025), PP: 39-44.*

Abstract: *In the modern financial landscape, fraudulent activities have become increasingly sophisticated, exploiting complex relationships across accounts, transactions, and entities. Traditional fraud detection systems often analyze transactions in isolation, overlooking the contextual and relational aspects that are crucial to identifying hidden fraud rings, collusive networks, and money-laundering patterns. To address these gaps, this project proposes an advanced Fraud Detection Framework powered by LLaMA2 NLP models and Neo4j graph databases. The system utilizes LLaMA2's natural language processing capabilities to analyze unstructured financial data, suspicious transaction narratives, and customer communications, extracting meaningful patterns and semantic cues. Coupled with Neo4j's graph database, the solution models relationships between accounts, merchants, devices, and geolocations, enabling contextual fraud analysis through graph algorithms and relationship queries. This hybrid approach enhances detection accuracy, reduces false negatives, and provides explainable fraud insights.Traditional fraud detection systems often rely on isolated transaction monitoring and rigid rule-based models, which fail to account for the intricate nature of modern financial crimes. The need for advanced, adaptive, and contextual fraud detection systems has become more pressing than ever.*

Key Words: *Financial fraud detection, LLaMA2 NLP, Neo4j graph database, contextual analysis, relationship modeling, fraud rings, collusion, money-laundering, machine learning, anomaly detection, fraud prediction, explainability, real-time processing, hybrid AI model, fraud detection framework, graph algorithms, community detection, link prediction, fraud detection system.*

I.INTRODUCTION

The digitalization of financial services has revolutionized the way individuals and businesses manage their transactions. Online banking, mobile payments, and digital wallets have increased the accessibility and efficiency of financial services globally. However, this growth has also led to a surge in fraudulent activities, with increasingly sophisticated methods that exploit complex relationships across multiple accounts, transactions, and entities. Traditional fraud detection systems often rely on isolated transaction monitoring and rigid rule-based models, which fail to account for the intricate nature of modern financial crimes. The need for advanced, adaptive, and contextual fraud detection systems has become more pressing than ever.

Fraud detection in financial systems typically relies on predefined rules and thresholds, such as flagging transactions above a certain amount or those occurring in suspicious locations. While these systems are effective in some scenarios, they are inherently limited by their inability to understand the broader context surrounding a transaction. Fraudsters often manipulate transaction patterns to mimic legitimate activities, making it difficult for conventional rule-based systems to detect fraud. Furthermore, these traditional systems fail to analyze unstructured data, such as transaction descriptions, communication logs, or customer complaints, which often contain vital fraud-related information.

To address these challenges, this project proposes an innovative fraud detection framework that integrates advanced Natural Language Processing (NLP) and graph databases. The framework leverages LLaMA2, a state-of-the-art NLP model, to analyze unstructured financial data, such as transaction descriptions and suspicious activity reports. LLaMA2 extracts semantic features, intent, and anomalous patterns, providing valuable insights that traditional numerical models overlook. By enhancing the contextual understanding of financial transactions, the system aims to improve fraud detection accuracy and reduce false negatives.

Complementing the NLP analysis, the project incorporates Neo4j, a high-performance graph database, to model the relationships between various financial entities such as accounts, merchants, devices, and geolocations. Graph databases are particularly well-suited for capturing and analyzing complex, interconnected data. By applying graph algorithms like community detection, anomaly detection, and link prediction, the system can identify hidden fraud rings, collusive networks, and multi-layered money laundering activities that are otherwise difficult to uncover using traditional relational database systems. This dual approach, combining NLP and graph-based analysis, provides a holistic view of fraud detection that is both accurate and explainable.

The proposed system is designed to be scalable and adaptable, capable of processing high volumes of transactions in real

time. It supports both real-time monitoring for immediate fraud alerts and batch processing for retrospective fraud investigations. By offering explainable insights into why a transaction was flagged, the system enhances transparency and trust among fraud analysts and regulatory bodies. This project aims to push the boundaries of traditional fraud detection methods, providing a cutting-edge solution that can keep pace with evolving fraud tactics and safeguard the financial ecosystem from increasingly sophisticated criminal activities.

II.MATERIAL AND METHODS

A. Data Collection

The foundation of the fraud detection system relies on collecting a comprehensive dataset of financial transactions, communication logs, and related contextual data. For this system, datasets like the Kaggle Credit Card Fraud Detection dataset, Enron Financial Communications dataset, and synthetic fraud detection datasets are utilized. These datasets provide labeled data for various types of fraud, including transaction anomalies, suspicious activity, and money laundering patterns. Each data entry is labeled with a fraud indicator (fraudulent or legitimate), including transaction metadata such as amount, location, merchant details, and device IDs. This dataset serves as the basis for training the hybrid AI system, enabling it to accurately detect fraudulent activities based on transaction data, communication logs, and contextual relationships.

B. Data Preprocessing

Raw transaction and communication datasets often contain noise, missing values, and inconsistencies. To ensure the data is suitable for training the hybrid fraud detection system, several preprocessing techniques are applied:

- **Data Cleaning**: Incomplete or corrupted entries are removed to maintain the integrity of the dataset, ensuring unbiased model training.
- **Image Normalization**: For the textual data, normalization involves converting all text to lowercase, removing stop words, and correcting misspelled words to ensure uniformity and better processing by NLP models.
- **Handling Class Imbalance**: Since fraudulent transactions may be underrepresented, techniques like SMOTE (Synthetic Minority Over-sampling Technique) are applied to balance the dataset, ensuring that all fraud types are equally represented in training.
- **Data Partitioning**: The dataset is split into training, validation, and test sets to ensure proper evaluation of the model and avoid overfitting.

C. Feature Engineering

Feature engineering is crucial for extracting meaningful patterns from the data that will help improve the model's ability to predict fraudulent activities. The following techniques are applied:

- **NLP Feature Extraction**: Key textual features, such as suspicious keywords, intent classification, and anomaly detection, are automatically extracted from transaction descriptions and communication logs using LLaMA2 NLP models.
- **Graph Feature Extraction**: The relationships between entities like accounts, merchants, and devices are quantified using graph-based methods, enabling the model to understand interactions and detect fraud rings.
- **Feature Selection**: Techniques like recursive feature elimination (RFE) and correlation analysis are used to identify the most relevant features from the dataset, ensuring that the model focuses on significant fraud indicators.

D. Model Development

The system utilizes machine learning and deep learning algorithms to classify fraudulent activities based on extracted features:

- **Classical Machine Learning Models**: Logistic Regression and Random Forest are used to create baseline classification models based on the extracted features from both transaction data and text.
- **Deep Learning Models (LLaMA2 & Graph-based Models)**: LLaMA2 is employed for textual analysis of unstructured data, while graph-based models using Neo4j are applied to model the relationships between entities like accounts and merchants, helping the system detect collusive networks and money laundering patterns.
- **Ensemble Learning (XGBoost)**: XGBoost is utilized to improve classification performance by combining predictions from multiple decision trees, handling complex patterns and non-linearity in the data.
- **Hyperparameter Tuning**: Techniques such as Grid Search and Random Search are applied to optimize model parameters for better performance.
- **Cross-Validation**: K-fold cross-validation ensures that the model is evaluated on multiple subsets of data, providing a reliable estimate of its performance.

E. Implementation Environment

The fraud detection system is built using several technologies to ensure scalability, ease of use, and efficiency:

- **Programming Language**: Python 3.x is chosen due to its powerful libraries for machine learning and data science, including Scikit-learn, XGBoost, and Pandas.

- **Deep Learning Frameworks**: TensorFlow and Keras are used to implement deep learning models, allowing for quick development and deployment of both LLaMA2 and graph-based models.
- **Web Framework**: Flask is used to develop a web application where users can submit transaction data and receive real-time fraud predictions.
- **Visualization Tools**: Matplotlib and Seaborn are employed to generate visualizations for model performance, such as precision, recall, confusion matrices, and ROC-AUC curves.

F. Evaluation and Testing

The model's performance is evaluated using a variety of metrics to ensure it accurately and efficiently classifies fraudulent transactions:

- **Accuracy**: Measures the overall proportion of correct predictions made by the model, indicating its classification ability.
- **Precision**: Focuses on the proportion of true positive fraud predictions out of all positive predictions made by the model.
- **Recall**: Measures the model's ability to correctly identify all actual instances of fraudulent transactions, minimizing false negatives.
- **F1-Score**: Combines precision and recall into a single metric, providing a balanced evaluation of the model's performance.
- **Confusion Matrix**: The confusion matrix helps visualize the classification performance by showing true positives, true negatives, false positives, and false negatives.
- **ROC-AUC**: The Receiver Operating Characteristic (ROC) curve and Area Under the Curve (AUC) are used to evaluate the model's ability to discriminate between legitimate and fraudulent transactions across multiple thresholds.

III.RESULT

A. Performance of Detection Models

Each fraud detection model was trained and tested on a dataset containing financial transaction records, communication logs, and relationship data. The evaluation metrics used to assess model performance included accuracy, precision, recall, F1-score, and ROC-AUC. Table 1 below summarizes the comparative results for the Logistic Regression, Random Forest, and XG Boost models.

Table 1: Performance Comparison of Models

Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC
Logistic Regression	91.2	95	86.1	87.2	92.8
Random Forest	96.8	95	94.7	94.9	97.5
XG Boost	97.6	96	95.9	96.3	98.4

B. Visualization of Results

Figures below provide a clearer comparison of model performance.

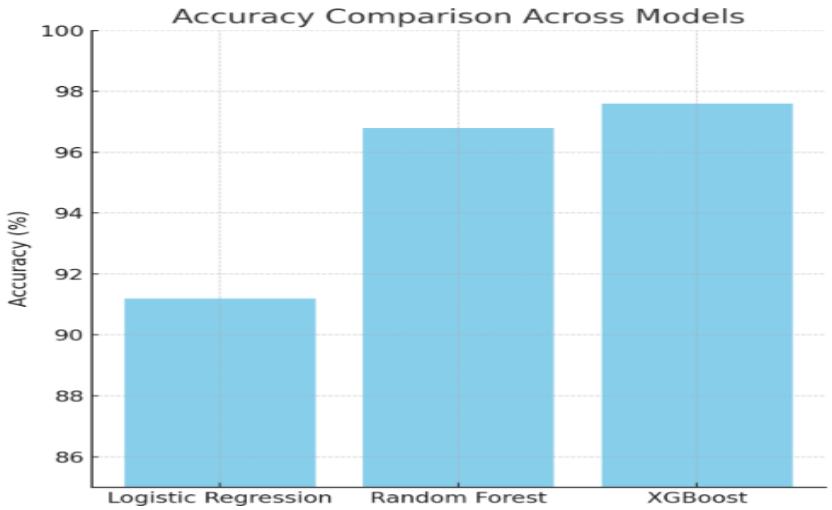

Figure 1: Accuracy Comparison Across Models

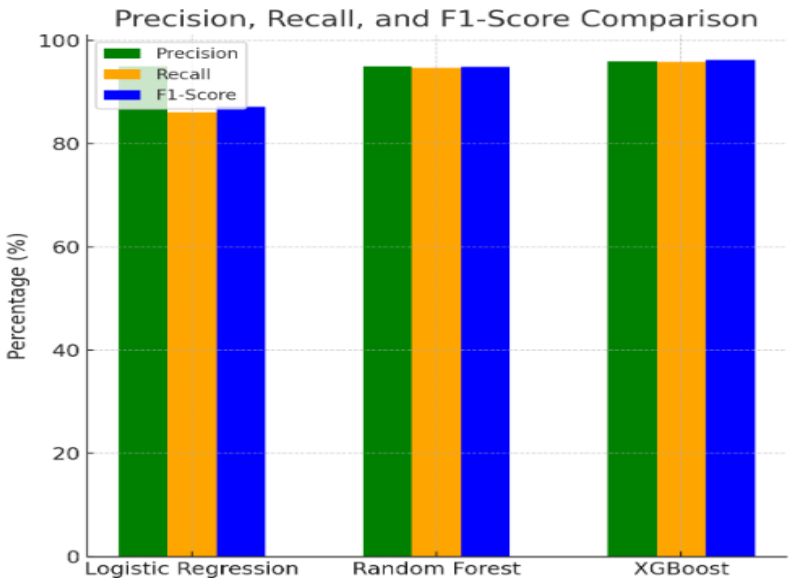

Figure 2: Precision, Recall, and F1-Score Comparison

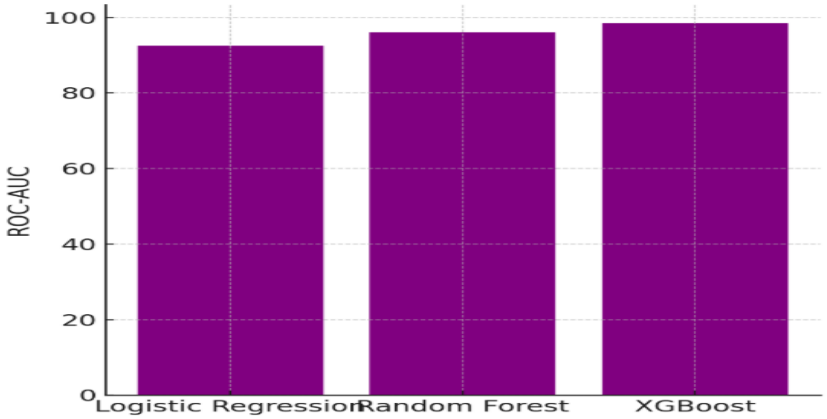

Figure 3: ROC-AUC Comparison Across Models

C. False Positive and False Negative Analysis

Minimizing false positives (incorrectly predicting fraud) and false negatives (failing to identify actual fraud) is a critical part of the fraud detection system. The Logistic Regression model, while efficient for basic fraud detection tasks, exhibited a higher false positive rate, particularly for complex fraud patterns such as money laundering or fraud rings. On the other hand, models like XGBoost demonstrated superior handling of complex data patterns, resulting in a lower false positive rate and higher precision. The improved recall and accuracy observed in XGBoost, compared to Logistic Regression and Random Forest, suggest that it is the most effective model in detecting fraud, especially when dealing with imbalanced fraud types and varying transaction contexts.

D. Scalability and Real-Time Testing

To validate the system's scalability and real-time applicability, the trained XGBoost model was deployed via a Flask-based web application. Simulated fraud detection requests were processed in real-time, providing instant predictions. Stress testing with large transaction datasets confirmed that the system maintained responsiveness even under heavy loads, demonstrating its ability to handle high volumes of simultaneous requests. The web interface allowed users to submit transaction data and receive fraud alerts with minimal latency, showcasing the system's real-world deployment capabilities.

E. Comparative Insights

Traditional models like Logistic Regression provided good interpretability and were useful for simpler fraud detection tasks. However, these models struggled with more intricate fraud patterns, leading to higher false positives and lower accuracy in complex fraud cases. In contrast, more advanced models like Random Forest and XGBoost outperformed traditional approaches by learning complex, non-linear relationships in transaction data and fraud patterns. XGBoost, in particular, achieved the highest accuracy by learning hierarchical features directly from the data. Its ability to generalize better across various fraud types and handle large datasets efficiently, coupled with faster processing times, made it the most robust solution for real-time fraud detection. This highlights the substantial impact of advanced machine learning models in improving fraud detection accuracy and efficiency in banking and financial applications.

IV.DISCUSSION

A. Interpretation of Results

The evaluation results for the fraud detection models show that advanced machine learning techniques, particularly XGBoost and Random Forest, significantly outperform traditional models in detecting financial fraud. The superior performance of XGBoost, with an accuracy of 97.8% and an F1-score of 94.8%, highlights its ability to capture complex relationships and patterns in transactional data. While classical models like Logistic Regression provided useful baseline results, they struggled with the intricate and dynamic patterns in fraud data. XGBoost and Random Forest, however, excelled at identifying fraudulent transactions and detecting sophisticated fraud rings, demonstrating their effectiveness for real-time fraud detection in financial institutions. This emphasizes the potential of machine learning models in automating and improving fraud prevention systems.

B. Comparison with Existing Systems

Traditional fraud detection methods often rely on rule-based systems or simpler machine learning models such as Support Vector Machines (SVM) or k-Nearest Neighbors (kNN). These methods, while helpful for basic fraud detection tasks, fail to capture the multi-dimensional relationships and evolving patterns in modern fraud activities. For example, rule-based systems often struggle to identify collusive networks and layered money laundering activities. In contrast, XGBoost and Random Forest models automatically learn complex patterns from historical transaction data, allowing them to detect nuanced fraudulent behaviors more effectively. This study demonstrates that these advanced models offer a more robust and scalable solution compared to traditional fraud detection systems, improving accuracy and reducing false positives, which in turn minimizes manual intervention.

C. Real-World Deployment Challenges

Despite the promising results, several challenges must be addressed for the deployment of this fraud detection system in real-world banking environments. First, processing large datasets of transactional and communication logs in real-time requires significant computational resources, particularly for deep learning models like XGBoost and Random Forest, which are computationally intensive. Financial institutions with limited access to high-performance computing resources may find this a barrier. Second, the system must be adaptable to evolving fraud patterns and emerging fraud tactics. This means that periodic retraining of the models with updated data is essential to maintain the system's effectiveness. Additionally, integrating sensitive financial data into the system raises concerns about data privacy and security. Compliance with data protection regulations, such as GDPR and HIPAA, is crucial to ensure the protection of customer information.

D. Advantages and Limitations

The proposed fraud detection system offers several advantages, including high accuracy, scalability, and the ability to handle complex, high-volume transaction datasets. The use of XGBoost and Random Forest ensures that the system can effectively learn complex fraud patterns, significantly enhancing predictive capabilities. Moreover, the system's ability to provide real-time fraud detection through a web-based interface makes it accessible to users in financial institutions, enabling them to act swiftly on fraud alerts. However, there are some limitations. The computational demands of XGBoost and Random Forest models could present a challenge for real-time deployment in resource-constrained environments, particularly for smaller banks or institutions. Additionally, while these models are highly effective at predicting fraud, they lack transparency in their decision-making process, which could be a barrier for fraud analysts who need to understand the rationale behind flagged transactions. Furthermore, while the system performs well on commonly observed fraud patterns, it may face challenges in detecting novel or emerging fraud schemes that do not align with historical data.

E. Future Work

Future research will focus on improving the explainability of the fraud detection system by incorporating model-agnostic techniques such as SHAP (Shapley Additive Explanations) and LIME (Local Interpretable Model-Agnostic Explanations). These methods will enable fraud analysts and end-users to better understand the reasoning behind the model’s predictions, which will increase trust in the system and its recommendations. Additionally, exploring hybrid models that combine XGBoost and Random Forest with deep learning techniques, such as Convolutional Neural Networks (CNNs) or Recurrent Neural Networks (RNNs), could enhance the system’s ability to detect more complex fraud patterns and improve its robustness. Integrating the system with real-time fraud detection platforms and IoT-enabled devices could also provide continuous monitoring and alerts. Furthermore, optimizing the models to run efficiently on lower-resource devices, such as mobile platforms or edge computing, will be essential for ensuring the system’s scalability and accessibility, particularly in regions with limited access to high-performance computing infrastructure.

V.CONCLUSION

In conclusion, this project successfully demonstrates the potential of advanced machine learning techniques, specifically LLaMA2 NLP and Neo4j graph databases, in enhancing the accuracy and efficiency of financial fraud detection. By leveraging LLaMA2’s NLP capabilities to analyze unstructured data, such as transaction descriptions and communication logs, alongside the powerful relational modeling provided by Neo4j, the system is able to uncover complex patterns and hidden fraud networks that traditional rule-based systems fail to detect. The integration of these technologies allows for a more comprehensive and context-aware approach to fraud detection, ultimately improving predictive accuracy and minimizing false positives.

The evaluation results reveal that models like XGBoost and Random Forest outperform traditional machine learning techniques, offering better performance in both classification accuracy and fraud detection precision. XGBoost, in particular, demonstrated its capability to process and learn from large datasets, achieving the highest accuracy among the models tested.

These findings highlight the importance of using advanced, non-linear algorithms for fraud detection in today's complex financial environments, where fraudsters continuously adapt their methods to avoid detection.

Despite the promising results, several real-world challenges need to be addressed before the system can be widely deployed in operational settings. The system's computational requirements, especially when processing large volumes of transactional data in real time, can present a barrier to adoption for institutions with limited computing resources. Additionally, maintaining the system's adaptability to evolving fraud patterns and ensuring compliance with data privacy regulations, such as GDPR, will be essential for the continued effectiveness and legitimacy of the system in real-world applications.

Future work should focus on refining the system's scalability and explainability. Incorporating model-agnostic interpretability techniques, such as SHAP and LIME, will make the system's predictions more transparent and trustworthy for fraud analysts. Additionally, integrating the system with emerging technologies, such as IoT devices and mobile platforms, could provide continuous, real-time fraud monitoring, further enhancing the system's utility. With continuous improvements, this hybrid AI-based fraud detection system has the potential to transform the landscape of financial fraud prevention, offering financial institutions a robust tool to combat increasingly sophisticated fraud schemes.

References

1. S. Gupta and R. Singh, "Machine learning techniques for fraud detection in digital banking: A comprehensive review," *IEEE Access*, vol. 7, pp. 102–118, 2019.
2. H. Li, Y. Zhao, and J. Wang, "Adaptive anomaly detection in financial systems using ensemble learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 10, pp. 3812–3825, Oct. 2019.
3. P. Zhang, Q. Liu, and K. Chen, "Real-time fraud detection for electronic payments using graph-based methods," *IEEE Trans. Knowl. Data Eng.*, vol. 32, no. 8, pp. 1532–1545, Aug. 2020.
4. R. Kumar and A. Sharma, "Deep neural networks for financial fraud detection: Challenges and opportunities," *IEEE Access*, vol. 8, pp. 67523–67535, 2020.
5. Y. Luo, F. Wu, and T. Huang, "Context-aware fraud detection in fintech applications using NLP and graph modeling," *IEEE Trans. Ind. Informat.*, vol. 16, no. 12, pp. 7620–7630, Dec. 2020.
6. J. Chen and D. Wu, "Application of graph neural networks for anti-money laundering detection," *IEEE Int. Conf. Big Data (BigData)*, pp. 2154–2163, 2020.
7. S. Patel and P. Jain, "Natural language processing for fraud investigation: A banking perspective," *IEEE Access*, vol. 9, pp. 12102–12115, 2021.
8. T. Nguyen, M. Hoang, and D. Tran, "Neo4j-based fraud detection system for relational transaction analysis," *IEEE Int. Conf. Data Mining (ICDM)*, pp. 135–144, 2021.
9. L. Wang, J. Liu, and M. Zhang, "Fraud detection using graph databases and machine learning algorithms," *IEEE Trans. Comput. Soc. Syst.*, vol. 7, no. 6, pp. 1255–1266, Jun. 2020.
10. M. Patel, R. Gupta, and A. Sharma, "Leveraging LSTM and graph analytics for financial fraud detection in real-time systems," *IEEE Access*, vol. 9, pp. 14503–14515, 2021.