International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 12 Issue: 04 | Apr 2025 www.irjet.net p-ISSN: 2395-0072

Machine Learning-Driven Detection of UPI Fraudulent Activities

Ayuti Khedekar[1], Prof. Ganesh Manerkar[2]

[1]*Student, Department of Information Technology and Engineering, Goa College of Engineering, Farmagudi, Goa, India*
[2]Assistant Professor, *Department of Information Technology and Engineering, Goa College of Engineering, Farmagudi, Goa, India*

Abstract - The rapid adoption of Unified Payments Interface (UPI) has revolutionized digital payments in India, offering seamless and instant transactions. However, with increased usage, fraudulent activities like phishing, vishing, fake payments, and QR code scams have surged. Traditional rule-based fraud detection systems struggle to address these sophisticated tactics. This paper proposes a machine learning-based solution that integrates advanced models, including Random Forest, XGBoost, LSTM, and CNN, to detect and classify various fraud types in real-time, enhancing UPI transaction security.

The system combines a hybrid approach that processes transaction data, QR code images, and audio from vishing attempts to identify fraud patterns. It is trained on a diverse dataset of labeled UPI transactions, QR codes, and voice data, enabling it to detect evolving fraud behaviors. The model's real-time alert system ensures timely detection, with promising results in accuracy. This approach provides a scalable solution to improve the security of digital payments, with future enhancements aimed at refining detection capabilities and adapting to new fraud tactics.

Key Words: **UPI, Fraud Detection, Machine Learning, Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM)**

1.INTRODUCTION

The Unified Payments Interface (UPI) has quickly become a key element of India's digital payment ecosystem, offering users fast, secure, and affordable ways to make financial transactions. As UPI's popularity has surged, it has revolutionized how individuals and businesses manage daily payments. However, this growth has also attracted fraudsters, leading to an increase in fraudulent activities such as phishing, vishing, fake payment confirmations, and QR code scams. These evolving fraud tactics pose significant challenges to the security and reliability of UPI, affecting both users and financial institutions.

This project seeks to address the growing issue of UPI fraud by implementing a machine learning-based fraud detection system. The system integrates multiple models, including Random Forest, XGBoost, Long Short-Term Memory (LSTM), and Convolutional Neural Networks (CNN), to detect and classify different types of fraud in UPI transactions. By analyzing factors like transaction amounts, timestamps, user behavior, QR code images, and audio data from vishing attempts, the system can adapt to emerging fraud patterns for suspicious activities.

The primary aim is to enhance the security of digital payments and foster trust by offering a fast, automated solution that accurately detects fraud with minimal human involvement. Traditional rule-based systems often fail to keep up with the constantly changing tactics of fraudsters, missing subtle anomalies in transaction patterns.

1.1 OBJECTIVE

The main goal of this project is to create an intelligent fraud detection system for UPI transactions that uses machine learning techniques to automatically detect and categorize different types of fraudulent activities. These include vishing (voice phishing), fake payments, impersonation, and counterfeit QR codes. The system is designed to provide real-time fraud detection and classification, ensuring that any suspicious transactions are quickly flagged for further investigation. By utilizing machine learning algorithms such as Random Forest, XGBoost, Long Short-Term Memory (LSTM), and Convolutional Neural Networks (CNN), the system will analyze transaction data, QR code images, and audio recordings to identify fraud with a high level of accuracy.

Another important objective of this project is to develop a scalable and adaptive fraud detection framework that can keep up with the ever-changing nature of fraudulent activities. As new fraud patterns emerge, the system will integrate incremental learning, enabling it to continuously update itself with new data without needing to be retrained from scratch. This flexibility ensures that the model remains effective in detecting evolving fraud tactics. Ultimately, the project aims to strengthen the security of UPI transactions, provide valuable insights to financial institutions, and support efforts to maintain trust in digital payment platforms.

1.2 METHODOLOGY

The methodology for detecting UPI fraud in this project utilizes a machine learning framework that processes various data sources to identify fraudulent activities such as fake payments, vishing, and QR code scams. The system is designed to handle different fraud types through a combination of specialized machine learning models, integrated into a unified structure that ensures real-time, accurate detection. Data for the fraud detection system is collected from diverse sources, including transaction records, audio data (for vishing detection), and QR code images. The raw data undergoes preprocessing to ensure consistency and quality, including cleaning, feature extraction, and normalization. This step ensures that the data is in the right format for model training, helping the system learn effectively.

Feature engineering plays a key role in identifying relevant patterns associated with fraudulent activities. By carefully selecting features that highlight specific characteristics of each fraud type, the system is able to enhance its ability to detect anomalies in transaction patterns, audio signals, and visual data from QR codes. These features are crucial in enabling the machine learning models to distinguish between legitimate and fraudulent activities with high precision.

A variety of machine learning algorithms are employed, each chosen based on the type of data it will process. These models are trained using historical datasets that include both fraudulent and legitimate data. The training process involves fine-tuning the models to identify complex fraud patterns in the input data. To optimize performance, hyperparameter tuning is conducted, adjusting key parameters such as learning rate and batch size to improve the accuracy of the predictions. Regularization techniques are applied to prevent overfitting, ensuring that the models generalize well to new, unseen data.

Once the models are trained, they are integrated into a cohesive system that can dynamically select the appropriate model based on the type of input data. This system is capable of processing data in real-time, immediately detecting fraud and providing alerts. Post-processing techniques are employed to refine model predictions, further reducing false positives and improving the system's accuracy. Additionally, the system incorporates incremental learning, allowing the models to adapt to new data and emerging fraud tactics, ensuring continuous improvement and relevance over time.

1.3 DATASET CHARACTERISTICS

The dataset used in this project is a diverse collection that captures various aspects of fraud detection in digital payment systems. It includes multiple data types such as transaction records, images, audio recordings, and text-based communication, each providing crucial insights into specific fraud activities. These diverse data sources work together to identify fraudulent behaviors across the payment ecosystem.

At the core of the dataset is transaction data, containing key features like transaction amounts, timestamps, user and recipient details, location, and merchant information. These features help identify financial and behavioural patterns, making it possible to detect anomalies that suggest fraudulent activity. To handle the imbalance between fraudulent and legitimate transactions, techniques such as data augmentation and resampling are used, ensuring the model learns effectively from both categories.

In addition to transaction data, the dataset incorporates images, audio, and text to detect fraud elements like tampered payment codes, fake identities, or communication-based fraud such as vishing. The combination of these various data sources enables the system to capture a comprehensive view of fraud activities, ensuring it remains adaptable and effective in real-world applications. As new fraud patterns emerge, the dataset is regularly updated, keeping the system relevant and enhancing its ability to detect complex, evolving fraud tactics.

2. RESULTS AND DISCUSSION

In this project, we developed a fraud detection system for UPI transactions by combining multiple data sources, including audio (for vishing), text (for impersonation), QR code images (for fake QR codes), and transaction data (for fake payments). Models were trained for each type of fraud, with algorithms tailored to process the specific data. The goal was to accurately classify fraudulent activities, using labeled datasets and evaluating the models based on key metrics such as accuracy, precision, recall, and F1 score.

The system utilizes several machine learning models: Random Forest Classifiers for vishing detection, Long Short-Term Memory (LSTM) networks, XGBoost for identifying fake payments, and Convolutional Neural Networks (CNN) for detecting fake QR codes.

Each model was assessed on its respective test data, with performance measured through standard classification metrics. In this section, we compare the models based on their accuracy, precision, recall, and F1 score, highlighting their effectiveness in detecting different fraud types.

The Vishing Model, created with the Random Forest Classifier, achieved an impressive accuracy of 97.88%, indicating its ability to reliably identify fraudulent and non-fraudulent calls. With a precision of 96.43% and recall of 99.39%, the model effectively minimizes false positives while accurately detecting fraud. The F1 score of 97.89%

ensures a balanced trade-off between precision and recall, further confirming its robust performance in distinguishing fraudulent calls.

These results show that the Random Forest Classifier is highly effective in detecting vishing within UPI fraud detection systems. The high precision and recall values suggest that the model is unlikely to misclassify legitimate calls as fraudulent, or vice versa, making it a reliable tool for identifying fraudulent calls.

	Precision	Recall	F1-Score	Support
0 (Not Fraud)	0.99	0.96	0.98	834
1 (Fraud)	0.96	0.99	0.98	816
Accuracy			0.98	1650
Macro avg	0.98	0.98	0.98	1650
Weighted avg	0.98	0.98	0.98	1650

Table-1: Classification Report of Random Forest

The Impersonation Detection Model was developed using a Long Short-Term Memory (LSTM) architecture, which is ideal for handling sequential data like text. After training the LSTM model, we achieved outstanding results, with the model showing exceptional performance across key metrics such as accuracy, precision, recall, and F1 score.

Performance Metrics:

- Accuracy: 0.9990
- Precision: 1.0000
- Recall: 0.9981
- F1 Score: 0.9991

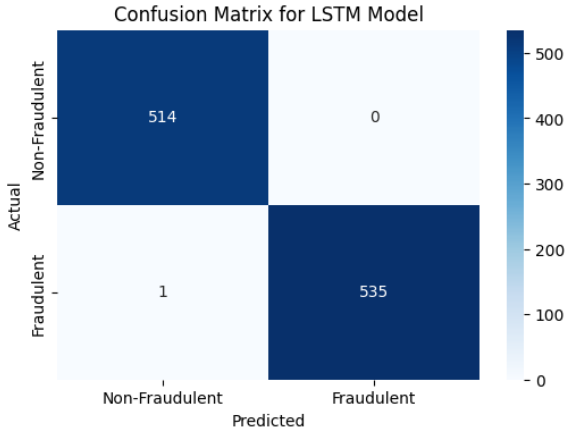

Fig -1: Confusion Matrix for LSTM Model

The Impersonation Detection Model, built using LSTM, showcases outstanding performance, as reflected in both the accuracy and loss plots.

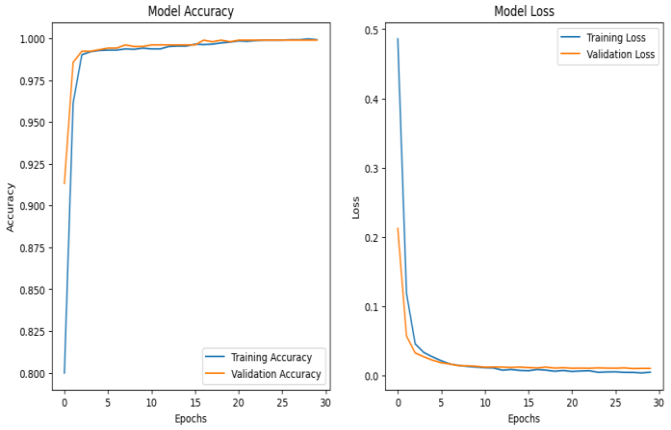

Fig -2: Plot Analysis for Impersonation Detection

The alignment of both training and validation accuracy curves indicates effective learning and generalization. The quick reduction in loss, followed by stabilization near zero, confirms minimal error and efficient learning, with no overfitting, showcasing the model's robustness. The Fake Payment Detection Model using XGBoost achieved an accuracy of 87.83% and an ROC AUC score of 0.9422, highlighting its strong performance in distinguishing fraudulent from legitimate transactions.

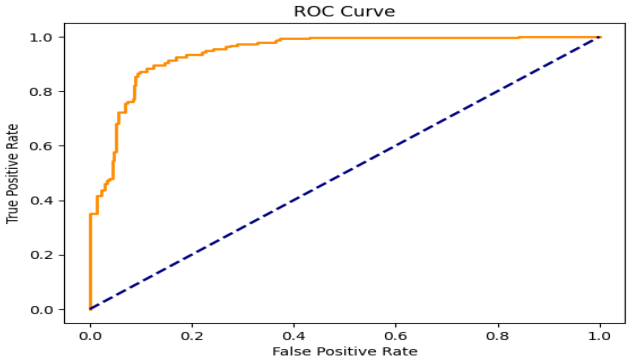

Fig-3: ROC curve for Fake Payment Detection

The ROC curve demonstrates the model's strong performance, showing high true positive rates and low false positive rates, confirming its effectiveness. The confusion matrix further supports this, indicating that the model correctly classified most transactions, with only a few false positives and negatives.

For Fake QR Code Detection, the CNN model showed significant improvement, steadily increasing accuracy throughout training and achieving near-perfect scores for both the training and validation sets, ensuring minimal overfitting.

The user-friendly UPI Fraud Detection System interface allows users to easily input transaction data and images for real-time fraud detection. It features an intuitive layout, enabling users to select parameters like language and location, and provides immediate fraud predictions based on the trained models.

3. CONCLUSIONS

The UPI Fraud Detection System, developed using machine learning approaches, offers an effective solution to detecting and addressing fraudulent activities in digital payment platforms. By integrating diverse data types—such as transaction details, audio data, text, and images—the system is equipped to handle a wide variety of fraud scenarios, including vishing, impersonation, fake payments, and QR code scams. The integration of models like Random Forest, LSTM, XGBoost, and CNN has resulted in a robust and precise system capable of accurately identifying different fraud types.

The system's performance, highlighted by high accuracy, solid ROC AUC scores, and minimal misclassification, underscores its effectiveness and reliability. Each model is tailored to process specific data types, enhancing the system's ability to handle complex fraud detection tasks. To ensure robust performance in real-world applications, techniques such as data augmentation and regularization have been applied, preventing overfitting and ensuring the models generalize well. The system also features an intuitive interface, allowing users and financial institutions to easily input data and receive real-time fraud predictions, making it accessible and practical for a wide range of users.

In the future, the system can be expanded with additional data sources, like biometric authentication and behavioural analytics, to improve accuracy and adaptability. As fraud tactics continue to evolve, ongoing updates and model improvements will ensure the system remains effective in combating emerging threats and securing digital payment ecosystems.

REFERENCES

[1] M. Naga Raju, Yarramreddy Chandrasena Reddy, Polavarapu Nagendra Babu, Venkata Sai Pavan Ravipati, Velpula Chaitanya. (2024). Detection of Fraudulent Activities in Unified Payments Interface using Machine Learning - LSTM Networks. 7th International Conference on Circuit Power and Computing Technologies (ICCPCT)

[2] N. Karthick, G. Leema Roselin, M. Tamilarasan, K. Kalaiselvi, J. Jayanthi (2024). Unified Payment Interface Imposture and Scam Detection Using Deep Learning and ANN. International Journal of Artificial Intelligence and Applications. Third International Conference on Smart Technologies and Systems for Next Generation Computing (ICSTSN) | 979-8-3503-9156 5/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICSTSN61422.2024.10671346

[3] Gangisetty Raj Charan, Dr.K Deepa Thilak (2023). Detection of Phishing Link and QR Code of UPI Transaction using Machine Learning. rd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA) | 979-8-3503-4363-2/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICIMIA60377.2023.10426613.

[4] Rishu, Amanveer Singh, Sarvesh Tanwar (2024). Revolutionizing Online Transaction Safety with CNN and GAN-Based Fraud Detection Strategies. Asia Pacific Conference on Innovation in Technology (APCIT) | 979-8-350361537/24/$31.0010.1109/APCIT62007.2024.10673599

[5] J. Kavitha, G. Indira, A. Anil kumar, A. Shrinita, D. Bappan (2024). Fraud Detection in UPI Transactions Using ML 2024 EPRA International Journal of Research and Development (IJRD) Volume: 9 | Issue: 4 | April 2024 Journal DOI: 10.36713/epra2016.

[6] B. Franklin Edburg1, K. Umadevi2, M. Vidya3 ,P. M. Ramesh Kumar(2022) Role of UPI Application Usage and Mitigation of Payment Transaction Frauds: An Empirical Study MDIM Journal of Management Review and Practice 2(1) 7 –22, 2024© The Author(s) 2024DOI: 10.1177/mjmrp.231222347 mbr.mjmrp.mdim.ac.inhttps://doi.org/10.1177/mjmrp.231222347.

[7] Deshpande, K., & Dam, L. B. (2021). Unified Payment Interface (UPI) platform: Conniving tool for Social Engineering Attack. Pacific Business Review International, 14(3), September 2021. Available at: https://www.researchgate.net/publication/357837725

[8] Muhammad Liman Gambo, Anazida Zainal, Mohamad Nizam Kassim "A Convolutional Neural Network Model for Credit Card Fraud Detection, "2022 International Conference on Data Science and Its Applications (ICoDSA) | 978-1-6654-8665-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICoDSA55874.2022.9862930.

[9] Mohammad Ziad Mizher, Ali Bou Nassif ''Deep CNN approach for Unbalanced Credit Card Fraud Detection Data," 2023 Advances in Science and Engineering Technology International Conferences (ASET) | 978-1-665454742/23/$31.0010.1109/ASET56582.2023.10180615

[10] S. K. Lokesh Naik, Ajmeera Kiran, Vadapally Praveen Kumar, Shanmukh Mannam, Yesarapu Kalyani, Manda Silparaj ,"Fraud Fighters - How AI and ML are Revolutionizing UPI Security," 2024 International Conference on Science Technology Engineering and Management (ICSTEM) | 979-8-3503-7691-3/24/$31.00©2024IEEE|DOI:.10.1109/ICSTEM61137.2024.10560740