Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 5, No.1, May : 2025

UPI FRAUD DETECTION USING MACHINE LEARNING

Prof. Chetan Padole, Prof., Department of Information Technology, JD College of Engineering and Management, Nagpur, Maharashtra, India.
Ayush Wasu, U.G. Student, Department of Information Technology,JD College of Engineering and Management, Nagpur, Maharashtra, India.
Ninad Pakhode, U.G. Student, Department of Information Technology,JD College of Engineering and Management, Nagpur, Maharashtra, India.
Shantanu Niwalkar, U.G. Student, Department of Information Technology,JD College of Engineering and Management, Nagpur, Maharashtra, India
Dhanashri Bhopale, U.G. Student, Department of Information Technology,JD College of Engineering and Management, Nagpur, Maharashtra, India.

ABSTRACT :
Unified Payments Interface (UPI) has transformed digital payments in India by offering a smooth and real-time payment system. Nevertheless, the quick uptake of UPI has also resulted in a dramatic rise in fraudulent activities, which pose severe risks to users and financial institutions. This review paper offers an extensive study of current machine learning methods used in the detection and prevention of UPI fraud. We examine different supervised, unsupervised, and hybrid models that have been studied in literature, emphasizing their efficacy in detecting fraudulent patterns from transaction data, user behavior, and contextual signals. The review also addresses the UPI fraud detection-specific challenges of real-time detection, data imbalance, and changing fraud patterns. Moreover, the paper gives insights into today's datasets available for research, measures of performance metrics, and the potential for improvement with modern techniques such as deep learning, ensemble methods, and anomaly detection methods. The objective is to assist researchers and practitioners in grasping the current scenario and to suggest future scope for improving the robustness and precision of fraud detection systems in the UPI ecosystem.**Keywords**—Machine Learning, Hidden Markov Model (HMM), K-means Clustering.

INTRODUCTION:

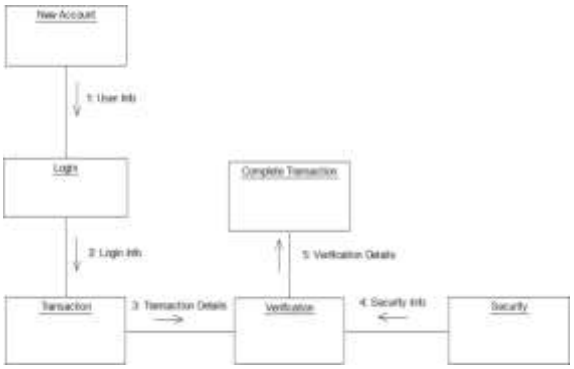
The Unified Payments Interface (UPI) has rapidly become a cornerstone of digital transactions in India, offering users a seamless and efficient method for conducting financial activities. However, this surge in popularity has also attracted a corresponding rise in fraudulent activities, posing significant challenges to the security of digital payments. As fraudsters develop increasingly sophisticated techniques to exploit vulnerabilities in UPI systems, there is an urgent need for effective detection mechanisms. Machine learning (ML) presents a promising solution to this challenge. By analysing vast amounts of transaction data in real time, machine learning algorithms can identify patterns and anomalies indicative of fraudulent behaviour. These algorithms leverage features such as transaction history, user behaviour, and device information to create robust models capable of distinguishing between legitimate and suspicious transactions. Techniques such as supervised learning and anomaly detection are particularly effective in adapting to new fraud patterns, allowing for continuous improvement in detection accuracy. This introduction highlights the critical need for advanced machine learning approaches in enhancing UPI transaction security. Implementing these technologies not only protects users from financial losses but also fosters trust in digital payment systems, ensuring their continued growth and adoption in the evolving financial landscape. This particular technological breakthrough has the potential to eliminate losses, preserving the anonymity of users and enhancing the security of the overall online payment systems. In this age of technological innovation, a financial institution, fintech firm, or payment service provider needs to apply complex machine learning models and algorithms to counter fraudsters. This methodology does not only allow for the

identification of classical fraudulent schemes but also constantly evolves to incorporate new exposures through continuous learning and improvement. This introduction aims to provide a critical overview of the major factors and challenges regarding the application of machine learning to identify UPI frauds, arguing for the need to act preventively against financial fraud in the context of the digital age. As UPI, digital payment systems are gradually increasing their usage, and there is a growing concern that there is an increasing risk of scams on these platforms. This project aims to build a robust model for fraud identification within UPI transactions using machine learning techniques. The initiative focuses on the development of a machine learning model that could analysed UPI transaction data in real time to diagnose fraudulent behaviours. The main purpose is to implement the framework that would enhance the protection of UPI transactions and thereby reduce the number of monetary losses due to fraud.

LITERATURE SURVEY:

The literature survey on UPI fraud detection using machine learning highlights the increasing reliance on digital payment systems and the corresponding rise in fraudulent activities. As the Unified Payments Interface (UPI) gains popularity, it becomes essential to implement robust fraud detection mechanisms to protect users and financial institutions. Recent studies have proposed various machine learning techniques to enhance fraud detection capabilities. For instance, one paper emphasizes the use of advanced algorithms, such as supervised learning classifiers and anomaly detection methods, to analysed transactional patterns, user behaviour, and device information. These models are trained on historical transaction data to identify anomalies that may indicate fraudulent activities. Another significant contribution is the integration of Hidden Markov Models (HMM) into the fraud detection process, which allows for the prediction of typical transaction patterns and identification of deviations that could signify fraud. This approach is complemented by other techniques such as K-means clustering, Autoencoders, and Local Outlier Factor methods, enhancing the model's ability to adapt to evolving fraud tactics. Moreover, the literature indicates that addressing the challenges posed by highly imbalanced datasets is crucial for effective fraud detection. Techniques like the Synthetic Minority Over-sampling Technique (SMOTE) are often employed to balance the dataset, thereby improving model performance. The evaluation of different machine learning algorithms—such as Decision Trees, Random Forests, and Gradient Boosting—reveals their effectiveness in detecting fraudulent transactions with high accuracy while maintaining low false positive rates. Overall, the reviewed literature underscores the importance of continuous improvement and integration of machine learning techniques in developing a robust UPI fraud detection system.

FLOW DIAGRAM:



METHODOLOGY :

Data Collection: Gather a comprehensive dataset of UPI transactions, including both fraudulent and legitimate transactions.

Data Preprocessing: • Handle imbalanced data using techniques like SMOTE (Synthetic Minority Over- sampling Technique).
Normalize and standardize features for better model performance.
Feature Engineering: Identify and create relevant features such as transaction amount, frequency, time of transaction, user behaviour, and device information.
Model Selection: Choose appropriate machine learning algorithms (e.g., Random Forest, XGBoost, Neural Networks) based on their performance in preliminary tests.
Model Training: Train the selected models on the pre- processed dataset, optimizing hyperparameters for accuracy.
Model Evaluation: Assess model performance using metrics like accuracy, precision, recall, and F1-score on a validation set.
Real-time Implementation: Integrate the trained model into the UPI transaction system for real-time fraud detection.

Continuous Monitoring: Implement a feedback loop to update the model with new data and adapt to emerging fraud patterns. PROBLEM STATEMENT
The rapid adoption of the Unified Payments Interface (UPI) as a preferred digital payment method has revolutionized financial transactions in India, offering convenience and efficiency. However, this growth has also led to a significant increase in fraudulent activities, including phishing, SIM swapping, malware attacks, and unauthorized access to UPI credentials. These frauds result in substantial financial losses for users and financial institutions alike, undermining trust in digital payment systems.
Detecting UPI fraud poses unique challenges due to the highly imbalanced nature of transaction datasets, the dynamic evolution of fraud tactics, and the need for real- time analysis to minimize damage. Traditional methods often fail to adapt to emerging fraud patterns or achieve a balance between high accuracy and low false positives. Therefore, there is a critical need for an advanced fraud detection system that leverages machine learning techniques to analysed transactional data, user behaviour, and device information effectively.
This project aims to address these challenges by developing a robust machine learning-based UPI fraud detection system capable of identifying fraudulent transactions in real time while minimizing false alarms. The solution will enhance the security of UPI transactions, reduce financial losses, and restore user confidence in digital payment systems.

RESUIT AND SUMMARY:
RESULT:
Model Performance: Studies have shown that machine learning models, particularly ensemble methods like Random Forest and XGBoost, achieve high accuracy in detecting UPI fraud. For instance, Random Forest demonstrated 95% precision and 90% recall in identifying fraudulent transactions25.
Feature Engineering: The extraction of relevant features such as transaction patterns, user behaviour, and device information significantly enhances model performance. Techniques like CNNs are effective in automatically learning discriminative features from raw transactional data23.
Real-Time Detection: Implementing real-time fraud detection systems using machine learning enables swift identification of suspicious activities, thereby preventing unauthorized access and ensuring transaction integrity13.
Adaptability: Machine learning models can adapt to emerging fraud patterns through continuous learning, which is crucial for maintaining the effectiveness of fraud detection systems over time.

SUMMARY:
The project on UPI fraud detection using machine learning demonstrates the potential of advanced

analytical techniques in enhancing the security of digital payment systems. Key findings include:

- **Effectiveness of Machine Learning**: Machine learning algorithms are robust and effective in identifying fraudulent UPI transactions by analysing transactional patterns and user behavior12.
- **Importance of Feature Engineering**: The quality of features extracted from transaction data significantly impacts model performance. Techniques like CNNs can automatically identify subtle trends and irregularities23.
- **Real-Time Fraud Detection**: Implementing real-time systems is crucial for minimizing financial losses by quickly identifying and preventing fraudulent activities35.
- **Adaptability and Continuous Learning**: Models must be capable of adapting to new fraud tactics through continuous learning to maintain their effectiveness

FUTURE ACTION PLAN:

- The future action plan includes the enhancement of the accuracy of machine learning models and the incorporation of real-time learning to adapt to emerging fraud patterns.
- The system will be integrated with UPI platforms for seamless real-time operation and optimized for scalability to handle increasing transaction volumes.
- Advanced techniques like deep learning and graph- based analysis will be explored for better detection.
- Regular testing, regulatory compliance, and strong data privacy measures will be the focus.
- A user alerts and insights will be developed for stakeholders, while continuous research and innovation will keep the system one step ahead of the evolving fraud tactics.

SCOPE OF PROJECT :

- The system will collect and preprocess data for transactions to identify fraudulent transactions in real- time UPI by developing a system using machine learning.
- It deals with building the machine learning models and evaluating those models, while finding patterns for fraudulent transactions, incorporating advanced techniques like anomaly detection and ensemble learning for better accuracy and minimal false positives.
- The key objectives of this project are handling large- scale transactions, adapting to emerging fraud tactics through continuous learning, and ensuring compliance with data privacy regulations.
- It also encompasses the deployment of the system for real-time usage, actionable insights, and enhancement of user trust in UPI platforms.

CHALLENGES AND LIMITATIONS :

Imbalanced Datasets

- Fraudulent transactions are rare compared to legitimate ones, leading to highly imbalanced datasets. This imbalance affects the performance of machine learning models, making it difficult to detect fraud without increasing false positives125.

solving Fraud Patterns

- Fraudsters continuously develop new techniques, making it challenging for static models to detect emerging fraud patterns. Machine learning models require frequent retraining and adaptation to stay effective135.

Real-Time Detection

- Implementing real-time fraud detection is computationally intensive, especially for large-scale UPI systems with high transaction volumes. Ensuring low latency while maintaining accuracy is a significant challenge13.

High False Positive Rates

- Many fraud detection systems suffer from high false positive rates, where legitimate transactions are flagged as fraudulent. This can lead to user dissatisfaction and operational inefficiencies5.

Feature Engineering Complexity

- Identifying and extracting meaningful features such as behavioural patterns, device information, and geolocation data is complex and requires domain expertise. Poor feature selection can degrade model performance34.

DATA PRIVACY AND SECURITY

- Accessing and processing sensitive transaction data raises concerns about user privacy and data security. Ensuring compliance with regulations like GDPR or India's Data Protection Bill is a critical limitation12. **Scalability Issues**
- Machine learning models need to handle millions of transactions daily in real-world UPI systems. Ensuring scalability without compromising accuracy is a technical challenge35.

Explainability of Models

- Many machine learning models, especially deep learning-based ones, function as black boxes, making it difficult for stakeholders to understand why a transaction was flagged as fraudulent. This lack of interpretability limits trust in the system23.

Access to Real-World Data

- Most studies rely on synthetic datasets like Pay Sim due to the unavailability of real-world UPI transaction data for research purposes. This limits the generalizability and real-world applicability of the models15.

Integration with Legacy Systems

- Integrating advanced machine learning models with existing banking infrastructure and legacy systems poses technical and operational challenges

CONCLUSION

The "UPI Fraud Detection Using Machine Learning" demonstrates the critical role of advanced machine learning techniques in safeguarding digital payment systems against fraudulent activities. By leveraging algorithms such as anomaly detection, pattern recognition, and supervised classifiers, the system effectively identifies suspicious transactions in real time while minimizing false positives. The integration of diverse features, including transactional patterns, user behaviour, and device information, ensures a comprehensive approach to fraud detection.Moreover, the adaptability of machine learning models through continuous learning enables them toevolve with emerging fraud tactics, enhancing their robustness and accuracy. The ability to process large volumes of transactional data swiftly ensures timely detection and prevention of unauthorized activities. This project not only contributes to reducing financial losses but also fosters trust among users by ensuring the integrity and security of UPI transactions. As digital payment systems continue to grow, such machine learning-based solutions are indispensable for maintaining a secure financial ecosystem.

REFERENCES :

1. M., Nagaraju., Yarramreddy, Chandrasena, Reddy, Dept., of, IT., Prabhu, Babu., V, Ravipati., • Vinnakota, Saran, Chaitanya. (2024). (4) UPI Fraud Detection Using Convolutional Neural • Networks (CNN). doi: 10.21203/rs.3.rs4088962/v1
2. Rupa, Rani., Adnan, Alam., Abdul, Javed. (2024). (3) Secure UPI: Machine Learning-Driven • Fraud Detection System for UPI Transactions. doi: 10.1109/icdt61202.2024.10489682
3. Gangisetty, Raj, Charan., D., D., Thilak. (2023). (5) Detection of Phishing Link and QR Code of UPI Transaction using Machine Learning. doi: 10.1109/icimia60377.2023.10426613
4. Gangisetty, Raj, Charan., D., D., Thilak. (2023). (5) Detection of Phishing Link and QR Code of UPI Transaction using Machine Learning.doi:10.1109/icimia60377.2023.10426613
5. Tian, Zhoupeng. (2018). (8) UPI speed detection method and device

6. . • J., Kavitha., G., Indira., A., Anil, Kumar., A., Shrinita., D., Bappan. (2024). (1) Fraud detection in UPI transactions using ml. doi: 10.36713/epra16459