International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 06 June 2024, pp: 98-100 www.ijaem.net ISSN: 2395-5252

UPI Fraud Detection Using Machine Learning

Shabreshwari R M, Shafiya Mehrooz, Sidra Fatima, Tanmai R B, Prof. Ganesh Manasali

Department of Electronics and Communication,
PDA College of Engineering, Kalabuaragi- 585103

Date of Submission: 28-05-2024 Date of Acceptance: 06-06-2024

ABSTRACT: Digital Fraud has become a threat across all the sectors. Its pivotal for any organization now to have a dedicated focus to detect and prevent fraud and increase their focus on Security. Machine learning algorithms have shown promise in analysing large volumes of transaction data to identify patterns and anomalies indicative of fraudulent transactions. Making use of a heterogeneous dataset that includes both authentic and fraudulent transactions, we utilize feature engineering and data preparation techniques to identify significant trends. Using past data, the chosen machine learning model is trained and its ability to distinguish between real and fraudulent transactions are evaluated. The model takes into account important features such transaction amount, timestamp, payer and payee data, location, and device information. For a thorough analysis, time-based features are given more weight. The model is included into the UPI system for real-time processing after the selected algorithm has been adjusted to maximize performance. In order to ensure prompt intervention, alert systems are implemented in tandem with the fraud detection model.

I. INTRODUCTION:

The landscape of financial transactions in India has undergone a remarkable transformation with the advent of digital payment systems, among which the Unified Payments Interface (UPI) stands as a hallmark of innovation and convenience. UPI, conceived and implemented by the National Payments Corporation of India (NPCI), has democratized financial inclusion by providing a seamless, interoperable, and instant platform for transferring funds between individuals, businesses, and institutions. In recent years, there has been a significant increase in the volume of financial transactions due to the expansion of financial institutions and the popularity of web-based e-commerce. Fraudulent transactions have become a growing problem in online banking, and fraud detection has always been challenging. Along with UPI development, the pattern of UPI fraud has always been updated. Fraudsters do their best to make it look legitimate, and UPI fraud has always been updated. They try to learn how fraud detection systems work and continue to stimulate these systems, making fraud detection more complicated. Therefore, researchers are constantly trying to find new ways or improve the performance of the existing methods. People who commit fraud usually use security, control, and monitoring weaknesses in commercial applications to achieve their goals. However, technology can be a tool to combat fraud. To prevent further possible fraud, it is important to detect the fraud right away after its occurrence.

Fraud detection in banking is considered a binary classification problem in which data is classified as legitimate or fraudulent. Because banking data is large in volume and with datasets containing a large amount of transaction data, manually reviewing and finding patterns for fraudulent transactions is either impossible or takes a long time. Therefore, machine learning-based algorithms play a pivotal role in fraud detection and prediction. Machine learning algorithms and high processing power increase the capability of handling large datasets and fraud detection in a more efficient manner. Machine learning algorithms and deep learning also provide fast and efficient solutions to real-time problems. In this paper, we propose an efficient approach for detecting UPI fraud that has been evaluated on publicly available datasets and has used optimised algorithms Random Forest, Light GBM, XGBoost, Decision treeand logistic regression individually. An ideal fraud detection system should detect more fraudulent cases, and the precision of detecting fraudulent cases should be high, i.e., all results

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 06 June 2024, pp: 98-100 www.ijaem.net ISSN: 2395-5252

should be correctly detected, which will lead to the trust of customers in the bank, and on the other hand, the bank will not suffer losses due to incorrect detection.Thus this project embarks on a quest to develop a robust and adaptive UPI Fraud Detection system capable of thwarting fraudulent activities and safeguarding users' financial assets.

II. OBJECTIVES:

1. To gather and pre-process a diverse UPI transaction dataset for fraud detection.
2. To identify pertinent features and create new metrics to enhance model discrimination.
3. To select and train machine learning algorithms for effective fraud detection.
4. To integrate the trained model into the UPI system for real-time processing and implement alert mechanisms.
5. To establish continuous monitoring, updates, and explore ensemble methods for model robustness.

III. METHODOLOGY:

TRAINING MODEL: Supervised machine learning is one of the category of machine learning where the model is trained by input data and expected output data. For creating such model, it is necessary to go through the following phases:

1. **MODEL CONSTRUCTION:** A model represents what was learned by a machine learningalgorithm. The model is the "thing" that is saved after running a machine learning algorithm on training data and represents the rules, numbers, and any other algorithm-specific data structures required to make predictions.
2. **MODEL TRAINING:** After model construction it is time for model training. In this phase, the model is trained using training data. At the end it will report the final accuracy of the model.
3. **MODEL TESTING:** During this phase a second set of data is loaded. This data set has never been seen by the model and therefore it's true accuracy will be verified.
4. **MODEL EVALUATION:** Model Evaluation is an integral part of the model development process. It helps to find the best model that represents our data and how well the chosen model will work in the future. Evaluating model performance with the data used for training is not acceptable in data science because it can easily generate overoptimistic and overfitted models. There are two methods of evaluating models in data science, Hold-Out and CrossValidation. To avoid overfitting, both methods use a test set (not seen by the model) to evaluate model performance.

IV. SYSTEM ARCHITECTURE:

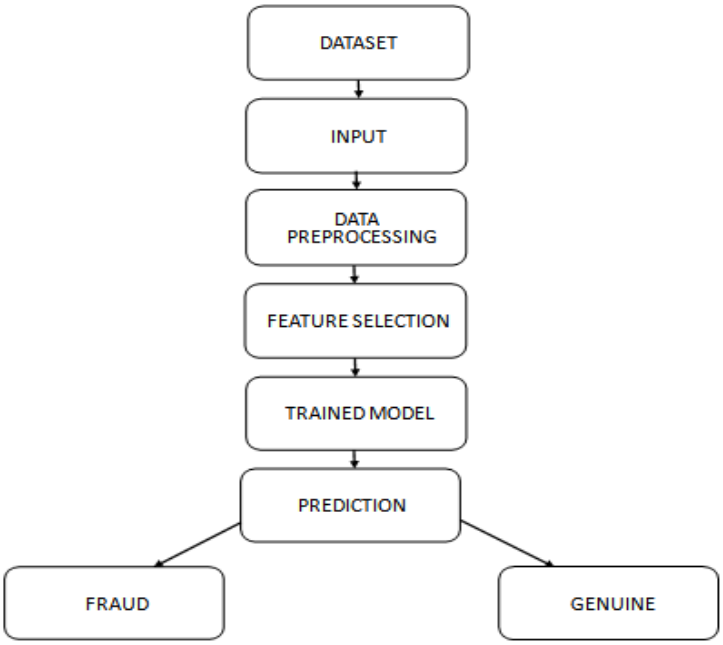

Fig 1: System Architecture of UPI fraud detection

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 06 June 2024, pp: 98-100 www.ijaem.net ISSN: 2395-5252

1. **Dataset**: Start with a labelled dataset containing information about UPI transaction with input parameters. Each transaction is labelled as genuine (Real) or fraudulent (fraud).
2. **Input**: Pre-process the dataset to ready it for input into the machine learning model. This involves handling missing data, encoding categorical variables, normalizing numerical features, and other necessary preprocessing steps.
3. **Feature Selection**: Identify relevant features that contribute to distinguishing between genuine and fraudulent transactions. Feature selection improves the model's performance by focusing on the most informative attributes, using techniques like correlation analysis or recursive feature elimination.
4. **Training the Model**: Split the dataset into training and testing sets. Train selected machine learning algorithms (Random Forest, XGBoost classifier, Logistic regression, Decision Tree, GBM) on the training set. During training, the model learns patterns that distinguish between genuine and fraudulent transactions.
5. **Evaluation**: Evaluate the performance of each model using the testing set. Common metrics for fraud detection, such as precision, recall, F1 score, and accuracy, are considered. Choose the algorithm with the highest accuracy for further use.
6. **Selection of Best Model**: Based on evaluation results, select the algorithm with the highest accuracy for fraud detection.
7. **Prediction:** Use the trained models to predict whether new, unseen transactions are genuine or fraudulent. Input the features of a new transaction into each model, and the model will output a prediction.

V. FUTURE SCOPE:

The future scope of UPI Fraud Detection System Using Machine learning has a lot of potential to get better. We're excited about trying out new and improved techniques in computer learning, using behaviour details. These improvements will not only make our system better but also let us work with people from around the world who have different skills.Looking ahead, we know it's important to keep up with the latest technology. As students working on this project, our main goal is to make our system work with the new changes in quantum computing while making sure it's clear and private. We also want to focus on making our project easy to understand by using a kind of AI that can explain how it works. By including these things in our future plans, we hope to keep our project working well in the always-changing world of technology.

VI. CONCLUSION:

In conclusion, developing and putting into action a UPI fraud detection system is a crucial step to make Unified Payments Interface transactions safer and more trustworthy. As digital payments become more widespread, it's essential to safeguard users and financial institutions from fraud. The proposed system, using advanced machine learning, aims not only to spot unusual transactions but also to adapt to new fraud patterns. By analysing past transaction data, the system intends to identify subtle signs of fraud, strengthening the overall security of UPI transactions. The method we plan to follow involves collecting and preparing data, creating a model, integrating it into the system, and continuously monitoring its performance. We will prioritize ethical considerations, user privacy, and compliance with regulations to ensure the responsible and legal use of the system. If successful, the UPI fraud detection system is expected to bring about better accuracy, real-time monitoring, adaptability to new threats, increased user trust, and reduced financial losses. The final output will include detailed documentation, user manuals, alert systems, and seamless integration with the existing UPI framework.

REFERENCES:

[1] Mahbuba Yesmin Turaba et al. “Fraud Detection During Financial Transactions Using Machine Learning and Deep Learning Techniques” in IEEE Oct 2022.
[2] Seyedeh Khadijeh Hashemi et al., “Fraud Detection in Banking Data by Machine Learning Techniques”, in IEEE Dec2022.
[3] Mr. Sunil S Mhamane and Mr. L.M.RJ Lobo “Internet Banking Fraud Detection Using HMM”, in IEEE July 2012.
[4] Pradheepan Raghavan and Neamat El Gayar “Fraud Detection using Machine Learning and Deep Learning”, in IEEE Feb 2020.
[5] G.Jaculine Priya and Dr.S.Saradha “Fraud Detection and Prevention Using Machine Learning Algorithms: A Review”, in IEEE 2021.