e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:07/Issue:01/January-2025 Impact Factor- 8.187 www.irjmets.com

FRAUDNET: UPI TRANSACTION FRAUD DETECTION USING MACHINE LEARNING

Prof. Dhananjay Subhash Mane*1, Mr. Prasad Vilas Gaikwad*2

*1Guide, Professor, Department Of Computer Science & Engineering, Vidya Vikas Pratisthan Institute Of Engineering And Technology, Solapur, Maharahtra, India.

*2Student, Department Of Computer Science & Engineering, Vidya Vikas Pratisthan Institute Of Engineering And Technology, Solapur, Maharahtra, India.

ABSTRACT

With the swift expansion of digital transactions, the Unified Payments Interface (UPI) has become a favored and convenient method for financial exchanges in today's world. Nevertheless, the growing dependence on digital platforms has also contributed to an uptick in fraudulent activities. This paper introduces a robust UPI fraud detection system that utilizes advanced machine learning techniques to bolster the security of digital transactions. The suggested system harnesses a wide array of features, including transaction patterns, user behavior, and device information, to develop a comprehensive fraud detection model. Machine learning algorithms, such as supervised learning classifiers and anomaly detection methods, are used to analyze historical transaction data and uncover patterns that indicate fraudulent activities.The model utilizes a labeled dataset comprising both legitimate and fraudulent transactions, enabling it to effectively differentiate between normal and suspicious activity. Keywords: Transaction, Payment, UPI, Attackers, Fraudulent, Hoaxers, Money, Dataset. Random forest; decision tree; logistic regression; machine learning; gradient boosting method; confusion matrix

I. INTRODUCTION

This technology has the potential to reduce financial losses, safeguard user privacy, and improve the overall security of digital payment systems. In a landscape characterized by rapid technological change, it is essential for financial institutions, fintech companies, and payment service providers to adopt sophisticated machine learning models and algorithms to outpace fraudsters. This strategy not only aids in identifying established fraud patterns but also adjusts to new threats through ongoing learning and refinement. This introduction will outline the key elements and challenges associated with UPI fraud detection utilizing machine learning, emphasizing the necessity of staying proactive in the continuous fight against financial fraud in today's digital world. As the usage of digital payment systems like UPI (Unified Payments Interface) grows, so does the concern regarding fraud on these platforms. This project seeks to create a strong fraud detection system for UPI transactions through the application of machine learning techniques. It concentrates on building a machine learning model capable of analyzing UPI transaction data in real-time to detect fraudulent activities. The main goal is to develop a system that improves the security of UPI transactions and minimizes financial losses caused by fraud.

II. LITERATURE SURVEY

In the realm of fraud detection, we frequently encounter datasets that are significantly imbalanced. In our analysis of the selected dataset (Paysim), we demonstrate that our proposed methods can effectively identify fraudulent transactions with remarkable accuracy and minimal false positives, particularly in relation to TRANSFER transactions. The challenge of fraud detection often presents a balancing act between accurately identifying fraudulent instances and avoiding the misclassification of legitimate transactions. This balance is typically a strategic decision that every digital payment company must confront. To address this challenge, we have introduced a class weight-based approach. Additionally, we can enhance our methodologies by utilizing algorithms such as decision trees to capitalize on the categorical features related to accounts and users within the Paysim dataset. The Paysim dataset can also be viewed as a time series, allowing us to exploit this aspect by developing time series-based models using algorithms like convolutional neural networks (CNN). In our current approach, we consider the entire set of transactions collectively for training our models.

We can make client particular models - which are based on user's past value-based behavior and utilize them to assist move forward our decision-making prepare. All of these, we accept, can be Exceptionally viable in moving

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:07/Issue:01/January-2025 Impact Factor- 8.187 www.irjmets.com

forward our classification quality on this dataset [1]. Now a days Advanced exchanges are quickly expanding as it comes about in expanding online Installment fakes as well. In reality, agreeing to the Reserve Bank of India, comparing Walk 2022 to Walk 2019, computerized installments have risen in volume and esteem by 216% and 10%, respectively. Individuals are beginning to go all-in with computerized exchanges, but one can't deny the security issues that linger, and know how when it comes to online installments. Few a long time prior, we might have scarcely seen the online installment, but nowadays UPI installment QR code introduced at doorstep.

This welcomed the hoaxers and aggressors to create false exchanges and trick individuals for a few sum of cash. Luckily, the online exchanges are observed and thus may be investigations utilizing the most recent instruments. In this framework, an endeavor is made to develop a machine learning demonstrate to recognize false exchanges in a transaction's dataset. [2] Fraud location for credit/debit card, credit defaulters and comparable sorts is achievable with the help of Machine Learning (ML) algorithms as they are well able of learning from past extortion patterns or verifiable information and spot them in current or future transactions. False cases are meager in the comparison of non-fraudulent perceptions, nearly in all the datasets. In such cases detecting false exchange are very troublesome.

The most successful way to pre-vent advance default is to distinguish non-performing loans as before long as conceivable. Machine learning calculations are coming into locate as proficient at dealing with such information with sufficient computing influence. In this paper, the rendering of distinctive machine learning calculations such as Choice Tree, Irregular Timberland, straight regression, and Angle Boosting strategy are compared for discovery and forecast of extortion cases utilizing credit false manifestations. Assist show exactness metric have been performed with disarray network and calculation of precision, accuracy, recall and F-1 score along with Recipient Working Characteristic (ROC) bends [3].

Monetary extortion, considered as beguiling strategies for picking up monetary benefits, has as of late ended up a broad danger in companies and organizations. Routine strategies such as manual confirmations and reviews are loose, expensive, and time consuming for recognizing such false exercises. With the approach of manufactured insights, machine-learning-based approaches can be utilized scholarly people to distinguish false exchanges by analyzing a expansive number of monetary information. Hence, this paper attempts to display a precise writing survey (SLR) that methodicallly surveys and synthesizes the existing writing on machine learning (ML)-based extortion location. Especially, the audit utilized the Kitchenhand approach, which employments well defined conventions to extricate and synthesize the significant articles; it at that point report the gotten comes about. Based on the indicated look strategies from well known electronic database libraries, a few thinks about have been accumulated. After inclusion/exclusion criteria, 93 articles were chosen, synthesized, and analyzed. The survey summarizes well known ML methods utilized for extortion discovery, the most popular extortion sort, and assessment measurements.

III. PROPOSED METHODOLOGY

Preparing Show: Supervised machine learning is one of the category of machine learning where the demonstrate is prepared by input data and anticipated yield information. For making such model,it is vital to go through the taking after stages:

1. Model construction: A model represents what was learned by a machine learning algorithm. The show is the "thing" that is spared after running a machine learning algorithm on preparing information and speaks to the rules, numbers, and any other algorithm specific information structures required to make predictions.

2. **Model preparing**: After model construction it is time for demonstrate preparing. In this stage, the demonstrate is prepared utilizing training data. At the conclusion it will report the last accuracy of the model.

3. **Model TESTING**: Amid this stage a second set of information is stacked. This information set has never been seen by the demonstrate and in this manner it's true precision will be confirmed.

4. **Model Evaluation** : Show Evaluation is an necessarily portion of the show development process. It makes a difference to discover the best demonstrate that represents our information and how well the chosen model will work in the future. Evaluating model execution with the information utilized for training is not worthy in information science because it can effortlessly create overoptimistic and overfitted models. There are two methods of assessing models in information science, Hold-Out and CrossValidation. To maintain a strategic

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:07/Issue:01/January-2025 Impact Factor- 8.187 www.irjmets.com

distance from overfitting, both methods utilize a test set (not seen by the model) to assess demonstrate execution.

IV. SYSTEM ARCHITECTURE

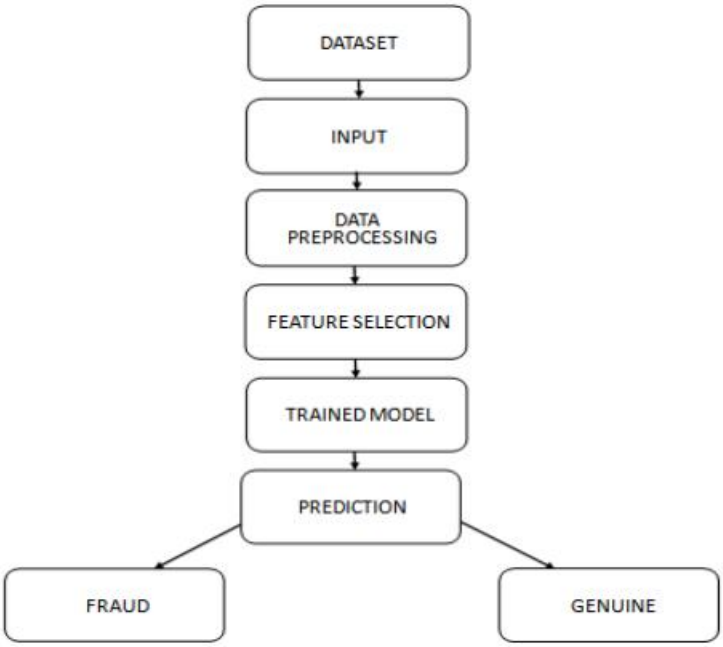

Figure 4.1: System architecture of FraudNet

1. Dataset: Begin with a named dataset containing data around UPI transaction with input parameters. Each exchange is labelled as veritable (Genuine) or fraudulent (fraud).

2. Input: Pre-process the dataset to prepared it for input into the machine learning show. This involves taking care of lost information, encoding categorical factors, normalizing numerical features, and other vital preprocessing steps.

3. Feature Choice: Recognize pertinent features that contribute to recognizing between genuine and false exchanges. Feature selection makes strides the model's execution by focusing on the most instructive attributes, using strategies like relationship investigation or recursive include elimination.

4. Preparing the Show: Part the dataset into training and testing sets. Prepare selected machine learning calculations (Irregular Forest, XGBoost classifier, Calculated regression, Decision Tree, GBM) on the preparing set. During preparing, the demonstrate learns designs that distinguish between veritable and fraudulent transactions.

5. Evaluation: Assess the execution of each model utilizing the testing set. Common metrics for extortion discovery, such as exactness, recall, F1 score, and precision, are considered. Choose the calculation with the most noteworthy precision for further use.

6. Determination of Best Demonstrate: Based on evaluation results, select the calculation with the highest accuracy for extortion detection.

7. Forecast: Utilize the prepared models to predict whether modern, concealed exchanges are genuine or false. Input the highlights of a new transaction into each demonstrate, and the demonstrate will output a expectation

V. CONCLUSION

The execution of a vigorous extortion location framework, fueled by machine learning strategies, is vital for enhancing the security of advanced exchanges and ensuring clients from monetary misfortunes. The fruitful implementation of this extend has the potential to essentially diminish the affect of UPI extortion, in this manner fostering client certainty in advanced installment frameworks. As innovation proceeds to progress, continuous endeavors in research and improvement are fundamental to remain ahead of developing dangers, eventually making a more secure and more secure advanced money related scene.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:07/Issue:01/January-2025 Impact Factor- 8.187 www.irjmets.com

VI. FUTURE SCOPE

Future improvements of the "UPI Extortion Discovery Utilizing Machine Learning" venture seem include consolidating more modern machine learning models such as Angle Boosting, XGBoost, and Neural Systems to improve the exactness and flexibility of extortion location. Furthermore, This development would cater to a more extensive audience, giving clients with a consistent and reliable involvement over different gadgets whereas upgrading the strength and viability of the extortion discovery framework.

ACKNOWLEDGEMENTS

I would like to express my sincere thanks to my project guide Prof. D.S.Mane , Department of Computer Science and Engineering, Vidya Vikas Pratishthan Institute of Engineering & Technology, Solapur for him systematic & stimulating suggestions and back-up helped me throughout the project work and writing of this synopsis report. It would not have been possible for me to complete this report within a short time without his encouragement, guidance and in time assessment.

I extend my sincere thanks to Principal Dr. A.N. Gaikwad and Prof. M.D. Katkar, Head of Department of Computer Science and Engineering, Vidya Vikas Pratishthan Institute of Engineering & Technology, Solapur for his kindly guidance in my synopsis report completion. I am thankful to my family for their constant support and inspiration. At the end I would like to thank all persons who directly or indirectly helped me to complete this synopsis report.

VII. REFERENCES

[1] Aditya Oza “Fraud Detection using Machine Learning” - https://github.com/aadityaoza/CS-229-project.

[2] Ms. Kishori Dhanaji Kadam, Ms. Mrunal Rajesh Omanna, Ms. Sakshi Sunil Neje, Ms. Shraddha Suresh Nandai. “Online Transactions Fraud Detection using Machine Learning” Volume 5, Issue 6 June 2023, pp: 545-548 www.ijaem.net

[3] M. Valavan and S. Rita “Predictive-Analysis-based Machine Learning Model for Fraud Detection withBoosting Classifiers” Computer Systems Science & Engineerin

[4] Abdulalem Ali 1,,Shukor Abd Razak 1,2,ORCID,Siti Hajar Othman 1ORCID,Taiseer Abdalla Elfadil Eisa 3,Arafat Al-Dhaqm 1,ORCID,Maged Nasser 4ORCID,Tusneem Elhassan 1,Hashim Elshafie 5 andAbdu Saif 6ORCID “Financial Fraud Detection Based on Machine Learning: A Systematic Literature Review”

[5] PayPal Inc. Quarterly results https://www.paypal.com/stories/us/paypalreports-third-quarter-2018-results

[6] A Model for Rule Based Fraud Detection in Telecommunications - Rajani, Padmavathamma - IJERT – 2012

[7] HTTP Attack detection using n−gram analysis - A. Oza, R.Low, M.Stamp - Computers and Security Journal - September 2014

[8] Scikit learn - machine learning library http://scikit-learn.org

[9] Paysim - Synthetic Financial Datasets for Fraud Detection
https://www.kaggle.com/ntnu-testimon/paysim1