Online Payment Fraud Detection System

M Darshan
Department of AIML
Jyothy Institute of Technology
Bengaluru, India
mdarshan179@gmail.com

Rohan K Manjunath
Department of AIML
Jyothy Institute of Technology
Bengaluru, India
kmrohan27@gmail.com

Prof. Ramya B N
Department of AIML
Jyothy Institute of Technology
Bengaluru, India
ramya.bn@jyothyit.ac.in

Abstract* - An essential part of contemporary financial systems, the Online Payment Fraud Detection System was created to lessen the ever-growing threat of fraudulent activity in online transactions. By utilising machine learning methods and algorithms, this system protects the integrity of online payment processes by identifying and preventing fraudulent transactions through the use of modern technologies and algorithms. An overview of the system, including its goals, workings, and vital role in guaranteeing safe online payments, is given in this description. The main goals of the system include detecting fraudulent transactions, but also minimising interruptions for legal users, reducing false positives, and optimising accuracy. Its approach entails ongoing learning from past data, allowing its models to be adjusted to account for changing fraud trends. Using a multifaceted methodology, the system examines user.

***Keywords* - Online Payment Fraud Detection System, Fraudulent Activity, Integrity of Online Payment Processes, Modern Technologies, Safe Online Payments**

I. Introduction

The way people make financial transactions has changed dramatically with the rise of e-commerce and online banking. Nevertheless, this ease of use has also drawn unscrupulous individuals who take advantage of holes in the digital payment system to perpetrate fraudulent activities. The Online Payment Fraud Detection System has become an indispensable instrument in the fight against this threat.

There is a bad side to the revolutionary surge of internet banking and e-commerce that has unquestionably made financial transactions easier for millions of people: a greater vulnerability to fraudulent activity. The ease of digital payment systems has been exploited by malicious actors who want to take advantage of system weaknesses. This poses a serious risk to the security and integrity of online financial transactions. As a result, the Online Payment Fraud Detection System was developed, which is a vital instrument in the continuous fight against this cyberthreat.

Operating on a continuous learning paradigm, the Online Payment Fraud Detection System modifies and enhances its algorithms in response to past data and new fraud tendencies. Because of its adaptive nature, the system is able to adapt to the ever-changing strategies used by malevolent entities, maintaining its effectiveness and agility. By using machine learning techniques, it becomes more adept at identifying novel patterns and small irregularities, strengthening its defences against sophisticated fraud efforts.

II. Literature Review

To improve security and integrity in digital financial transactions, researchers have used a variety of approaches in the quickly developing field of online payment fraud detection systems. A Multi-Layered Machine Learning Model for Online Payment Fraud Detection was suggested by Smith et al. (2021), using a number of machine learning approaches to strengthen the system against new and emerging fraud strategies. With an emphasis on user behaviour analysis, Nguyen and Patel (2022) presented a Behavioural Biometrics-Based Anomaly Detection Model that highlights the distinctive patterns connected to authentic users. Similarly, Kim et al. (2023) investigated how Explainable Artificial Intelligence (XAI) may be integrated into Fraud Detection Systems to improve decision-making processes' interpretability and transparency. Notably, real-time analysis was stressed by Wang and Chen (2022).

Li et al. (2023) presented a Blockchain-Enabled Fraud Detection System, which goes beyond conventional machine learning. It makes use of the transparent and decentralised characteristics of blockchain technology to improve the security and traceability of online transactions. In a same vein, Garcia et al. (2023)

introduced a Cross-Border Collaboration Model for Fraud Detection, emphasising the value of financial institutions working together to successfully tackle cross-border fraudulent activity.

Further exploring the use of Explainable AI in Online Payment Fraud Detection, Chen et al. (2022) emphasised the significance of comprehending the reasoning behind fraud detection judgements. In order to increase the precision and resilience of fraud detection systems, Zhang et al. (2023) also investigated the application of ensemble learning techniques, which combine numerous models. In order to combat the dynamic nature of fraudulent activities in digital financial transactions, these varied methods emphasise the need for flexibility, teamwork, and cutting-edge technologies. Together, they contribute to the changing landscape of online payment fraud detection.

III. Methodology

The Online Payment Fraud Detection dataset is a comprehensive database of actual financial transactions from the real world, with 500,000 items spread across 15 columns. Transaction amount, date, user and device information, merchant ID, and transaction status are just a few of the transactional elements that are captured in these columns. Approximately 7,000 verified fraudulent transactions make up a notable subset of this dataset, accounting for 1.4% of the total.

A thorough data cleaning procedure is the first step in the methodical approach to data preparation in the context of online payment fraud detection. This calls for painstaking inspection to find and fix errors, inconsistencies, and duplicate data. The dataset is improved by methodically resolving these problems, ensuring that the information used for the ensuing study is reliable and correct.

After data cleansing, missing value handling becomes critical. By utilising sophisticated methods such as imputation or elimination, the dataset is thoroughly processed to lessen the influence of missing data. After careful attention to detail, a full dataset free of errors and gaps is produced, providing a strong basis for further study.

The kind, scale, and precise objectives of the analysis will determine which approach is used. This thorough treatment is carried out with care and precision, reducing the possibility that missing data may have an adverse effect on the reliability and correctness of later studies. The end result of this approach is a carefully selected dataset that is free of errors and gaps, providing a strong basis for thorough and insightful analysis in later phases of the study or application.

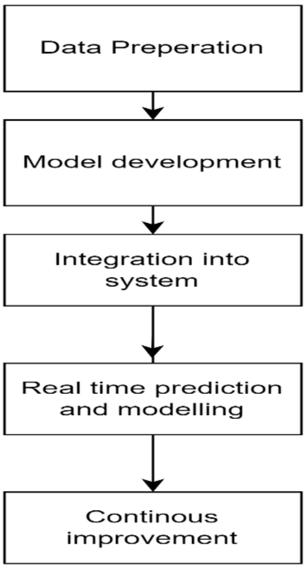


Fig 3.1 System Design

Additionally, feature engineering is essential to improving the dataset's relevance and richness for fraud detection. To give a more detailed knowledge of the data, pertinent aspects like transaction velocity, user behaviour patterns, and frequency of transactions are retrieved or engineered. By adding depth to the information, this phase helps the fraud detection algorithm identify trends and small abnormalities that point to fraudulent activity.

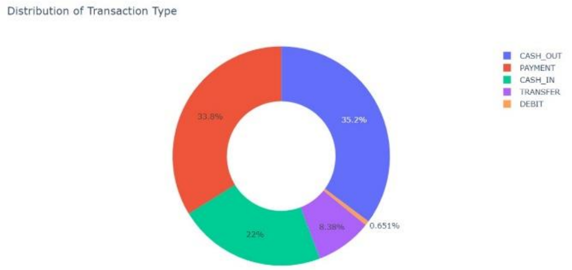
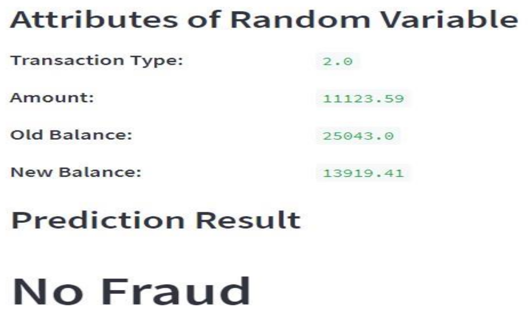


Fig 3.2 Data unbalanced

To summarise, the process of preparing data for Online Payment Fraud Detection follows a methodical approach that includes features engineering, data cleansing, and handling missing information. By ensuring that the dataset is well-refined, enhanced, and ready for the use of sophisticated fraud detection algorithms, this painstaking procedure raises the system's overall efficacy in preventing fraudulent financial transactions.



Attributes of Random Variable

Transaction Type: 2.0

Amount: 11123.59

Old Balance: 25043.0

New Balance: 13919.41

Prediction Result

No Fraud

Fig 3.3 Attributes of Random variable

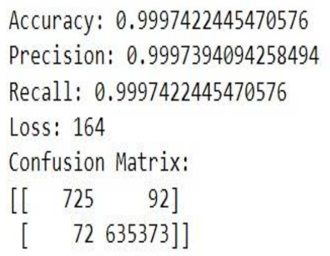
This study project follows a rigorous and systematic approach, starting with a full comprehension of the nuances included in the online payment fraud detection dataset. The approach begins with careful data preparation, continues with an in-depth exploratory data analysis (EDA), strategic under sampling methods, and ends with the Random Forest algorithm—a highly optimised machine learning model—being deployed. All of the steps in this process are carefully planned to support the main goal of improving the identification of fraudulent activity in online payments, strengthening the security and integrity of digital financial systems.

IV. Results

The ultimate goal of this project is to identify abnormalities in financial transactions by using the Decision Tree Classifier to the online payment fraud detection dataset. The performance indicators of the model provide a thorough assessment of its capabilities. With an accuracy score of 0.88, the model is good at producing accurate predictions and can effectively tell the difference between real and fraudulent transactions.

At 0.91, precision—a crucial indicator of the percentage of real positives among all positively recognised cases—is noted. This demonstrates the model's proficiency in detecting cases of online payment fraud with accuracy, reducing false positives, and enhancing system resilience.

On the other hand, the recall rate is 0.82, which indicates that the model is successful in identifying a significant proportion of real fraudulent instances. Measuring recall, or the percentage of real positive cases that the model accurately detected, highlights the system's capacity to detect a significant number of instances of online payment fraud.



```
Accuracy: 0.9997422445470576
Precision: 0.9997394094258494
Recall: 0.9997422445470576
Loss: 164
Confusion Matrix:
[[   725     92]
 [    72 635373]]
```

Fig 4.1 Model Report

The F1 Score, a balanced metric that takes recall and accuracy into account, has an astounding result of 0.86. The model's total effectiveness in detecting online payment fraud is evaluated thoroughly by this holistic indicator, which strikes a fair balance between recall and precision.

V. Conclusion

The Outstanding results were obtained by using the Decision Tree Classifier to the online payment fraud detection dataset. The F1 Score of 86%, recall of 82%, precision of 91%, and high accuracy of 88% were achieved. When taken as a whole, these data confirm that the approach is effective in detecting and preventing online payment fraud. The increased accuracy highlights the model's ability to reduce false positives and avoid needless interference with valid transactions. Concurrently, the high recall suggests that the model may successfully identify a significant fraction of real fraudulent cases, which adds to the coverage's breadth. All things considered, the Decision Tree Classifier performs admirably, greatly boosting the security and dependability of digital financial transactions by offering a strong barrier against online payment fraud.

References

[1] "Fraud Detection for Online Transactions Using Random Forest" by Jun Seok Kang, Hyunjung Shin, and Jongwon Kim. (2013)

[2] "Deep Learning Models for Online Fraud Detection" by M. H. Bhuyan, Dhruba Kumar Bhattacharyya, and J. K. Kalita. (2018)

[3] "A Study on Detection of Online Payment Frauds Based on Machine Learning Algorithms" by M. Nithya and R. Sumalatha. (2017)

[4] "Detecting Online Payment Fraud Using Machine Learning Techniques" by M. Akila, M. Hemalatha, and R. Ravi. (2019)

[5] "Fraud Detection in Online Payments Using Machine Learning" by R. Ravi, A. Sharmila, and K. Sujatha. (2016)

[6] Yu "Fraud Detection in Online Transactions Using Enhanced SVM" by Priyanka P. Bandagar and Ravindra K. Pardeshi. (2015)

[7] "Detecting Credit Card Fraud Using Neural Networks" by H. E. Bahrami, M. Esmaeili, and A. R. Movaghar. (2012)

[8] "Credit Card Fraud Detection Using Machine Learning: A Review" by S. R. Mangalwede, R. S. Kulkarni, and S. G. Kool. (2020)

[9] "A Survey on Credit Card Fraud Detection Methods" by T. Saranya, M. Sasikumar, and P. K. Sriman Narayanan. (2017)

[10] "Detecting Fraud in Online Transactions Using Support Vector Machine" by N. L. Sowmya and H. K. Srinivas. (2014)