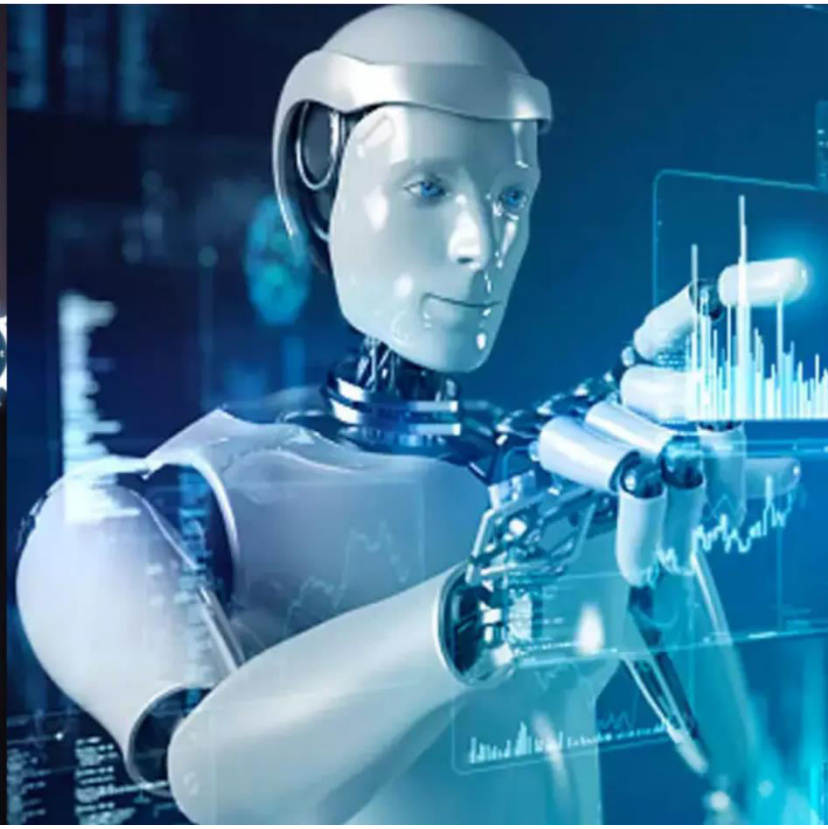
IJIRSET
International Journal of Innovative Research in Science | Engineering | Technology

ISSN(O): 2319-8753
ISSN(P): 2347-6710

International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 14, Issue 5, May 2025

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 5, May 2025

|DOI: 10.15680/IJIRSET.2025.1405503|

Online Payment Fraud Detection

R Ponmayil, Parnika S

Assistant Professor, Department of MCA, Vivekanandha Institute of Information and Management Studies, Tiruchengode, Tamil Nadu, India

Master of Computer Applications, Vivekanandha Institute of Information and Management Studies, Tiruchengode, Tamil Nadu, India

ABSTRACT: With the rapid growth of online transactions, the risk of fraudulent activities has significantly increased, posing serious threats to financial security. Traditional rule-based systems are often inadequate in detecting complex and evolving fraud patterns. This project, "Online Payment Fraud Detection using Machine Learning," aims to develop an intelligent system that can efficiently identify fraudulent transactions in real time. By analyzing historical transaction data, the system employs supervised machine learning algorithms such as Logistic Regression, Decision Trees, Random Forest, and XGBoost to classify transactions as legitimate or fraudulent.The model is trained on key features like transaction amount, time, location, and user behavior patterns. Data preprocessing techniques such as handling class imbalance with SMOTE and feature scaling are applied to enhance model performance.

KEYWORDS: Online Payment, Fraud Detection, Machine Learning, Classification Algorithms, Transaction Analysis, Anomaly Detection.

I. INTRODUCTION

Online payments have become an integral part of everyday transactions due to their convenience and speed. However, this rapid digitalization has also led to a significant increase in online fraud, posing major threats to individuals, businesses, and financial institutions. Traditional fraud detection methods, which rely on manually defined rules, often fall short in identifying new and sophisticated fraudulent behaviors. This has created a need for intelligent, automated systems that can detect and prevent fraudulent activities in real time.Machine learning offers a powerful solution by analyzing large volumes of transaction data and identifying hidden patterns that distinguish fraudulent transactions from legitimate ones. By using classification algorithms such as Logistic Regression, Decision Trees, Random Forest, and XGBoost, this project aims to build an effective fraud detection model that can adapt to changing fraud tactics. The system leverages key transaction features like amount, time, user behavior, and location to make accurate predictions.

II. OBJECTIVES

The main objective of this project is to develop an intelligent system that can effectively detect fraudulent online payment transactions using machine learning techniques. The project aims to analyze transaction data to identify patterns and behaviors that commonly indicate fraud. It involves preprocessing the dataset to handle issues such as missing values, class imbalance, and feature scaling, ensuring the data is suitable for training robust models. Various machine learning classification algorithms such as Logistic Regression, Decision Tree, Random Forest, and XG Boost will be implemented and compared to determine the most effective approach. The goal is to build a model that accurately classifies transactions as either fraudulent or legitimate, based on key features like transaction amount, time, location, and user behavior. The system will be evaluated using performance metrics including accuracy, precision, recall, F1-score, and AUC-ROC. Ultimately, the project seeks to create a scalable and real-time fraud detection solution that enhances the security and reliability of online payment systems.

III. LITERATURE REVIEW

Online payment fraud has emerged as a major challenge in the digital financial ecosystem, prompting extensive research in the field of fraud detection using machine learning. Traditional rule-based systems, although once effective, have proven inadequate in dealing with the dynamic and evolving nature of fraudulent behavior. As a result, researchers have

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 5, May 2025

|DOI: 10.15680/IJIRSET.2025.1405503|

turned to data-driven and intelligent approaches to enhance detection accuracy and efficiency.Bhattacharyya et al. (2011) proposed a credit card fraud detection system using machine learning techniques like decision trees and support vector machines (SVM), demonstrating that these methods can outperform conventional approaches. Their research highlighted the importance of feature selection and class imbalance handling, which significantly affect model performance. Dal Pozzolo et al. (2015) addressed the issue of class imbalance—common in fraud datasets—by applying resampling techniques such as SMOTE (Synthetic Minority Oversampling Technique) to improve classifier performance, particularly recall and precision.Sahin and Duman (2011) developed a cost-sensitive decision tree-based approach for fraud detection, emphasizing the need to reduce false positives while ensuring high fraud capture rates. Another study by Jurgovsky et al. (2018) utilized recurrent neural networks (RNNs) to capture temporal patterns in transaction sequences, showing that deep learning methods can further improve detection in complex, sequential datasets.Recent advancements also explore ensemble methods like Random Forest and XGBoost due to their robustness and ability to handle large, high-dimensional datasets. These algorithms have shown promising results in real-time fraud detection applications, offering a balance between speed and accuracy. Moreover, integrating unsupervised learning techniques for anomaly detection has gained popularity, especially when labeled data is scarce.

IV. EXISTING SYSTEM

The existing fraud detection systems in online payment platforms primarily rely on traditional rule-based approaches. These systems use predefined rules and thresholds, such as unusual spending amounts, login from unfamiliar IP addresses, or rapid transactions in a short period, to flag suspicious activities. While these systems can detect known fraud patterns, they often struggle with complex and evolving fraud techniques. Some systems also use basic statistical methods to analyze transaction behavior over time.

DISADVANAGES

- Not adaptive – Cannot detect new or evolving fraud patterns.
- High false positives – Often flags legitimate transactions as fraud.
- Manual process – Time-consuming and not suitable for real-time detection.
- Hard to update – Rule changes require manual intervention.Scalability Issues – Not designed for large-scale fleet operations.
- Poor on imbalanced data – Fails to detect rare fraud cases accurately.

V. PROPOSED SYSTEM

The proposed system utilizes machine learning algorithms to detect fraudulent online payment transactions in real-time. Unlike traditional rule-based systems, this system is data-driven and can automatically learn patterns from historical transaction data. It uses classification algorithms such as Logistic Regression, Decision Tree, Random Forest, and XGBoost to analyze transaction features (like amount, time, and user behavior) and accurately classify them as fraudulent or legitimate. The system also incorporates techniques like SMOTE to handle class imbalance and feature scaling to improve model accuracy.

ADVANTAGES:

- Adaptability – Learns from data and can detect new, unseen fraud patterns.
- High Accuracy – Provides better prediction performance using advanced algorithms.
- Real-time Detection – Can identify frauds instantly during the transaction process.
- Reduced False Positives – More accurate classification means fewer legitimate transactions are flagged.
- Handles Imbalanced Data – Uses techniques like SMOTE to improve detection of rare fraud cases.

VI. METHODOLOGIES

In this methodology for the Online Payment Fraud Detection system begins with collecting transaction data, which includes details such as transaction amount, time, user information, and fraud labels. The raw data undergoes preprocessing where missing values are handled, features are engineered to capture relevant transaction behaviors, and

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 5, May 2025

|DOI: 10.15680/IJIRSET.2025.1405503|

class imbalance is addressed using techniques like SMOTE to ensure the model can effectively detect rare fraudulent cases. Feature scaling is applied to normalize the data for better algorithm performance. Exploratory data analysis is performed to understand patterns and select important features. Several machine learning algorithms, including Logistic Regression, Decision Tree, Random Forest, and XGBoost, are then trained and fine-tuned on the prepared data. The models are evaluated using metrics such as accuracy, precision, recall, F1-score, and AUC-ROC, focusing on minimizing false positives and false negatives. The best-performing model is deployed to monitor transactions in real-time, flagging suspicious activities instantly. To maintain effectiveness, the system is designed to continuously learn and update its model with new transaction data, allowing it to adapt to evolving fraud tactics and improve detection accuracy over time.

MODULES

- Data Collection Module
- Data Preprocessing Module
- Exploratory Data Analysis (EDA) Module
- Model Training Module
- Model Evaluation Module
- Fraud Detection Module
- Model Updating Module

MODULE DESCRIPTION

Data Collection Module
Collects and aggregates transaction data from various sources, including payment gateways and financial databases, which contains details like transaction amount, time, user info, and fraud labels.

Data Preprocessing Module
Cleans the data by handling missing values and duplicates, performs feature engineering to create useful attributes, applies techniques like SMOTE to balance the dataset, and scales features for better model training.

Exploratory Data Analysis (EDA) Module
Analyzes data patterns and relationships among features to identify key indicators of fraud and select the most relevant features for modeling.

Model Training Module
Trains multiple machine learning algorithms such as Logistic Regression, Decision Tree, Random Forest, and XG Boost on the processed data, and tunes hyper parameters to optimize performance.

Model Evaluation Module
Assesses the performance of each model using metrics like accuracy, precision, recall, F1-score, and AUC-ROC to select the best-performing model for fraud detection.

Fraud Detection Module
Deploys the selected model to classify new incoming transactions in real time, flagging suspicious activities for further investigation.

Model Updating Module
Continuously updates the fraud detection model by retraining it with new transaction data to adapt to evolving fraud patterns and maintain detection accuracy.

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 5, May 2025

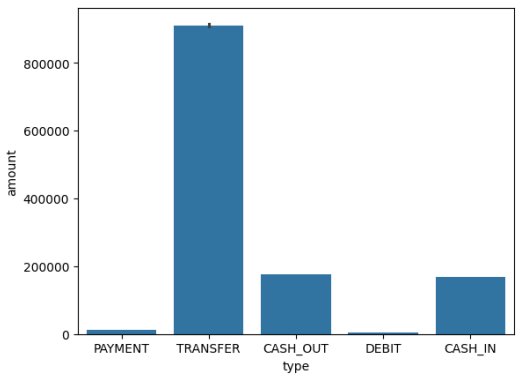
|DOI: 10.15680/IJIRSET.2025.1405503|

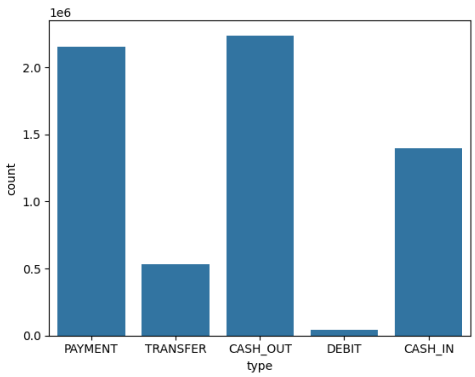
VII. EXPERIMENTAL RESULT

Figures shows the results (a) Info (b) Describe Data (c) Barplots

	step	type	amount	nameOrig	oldbalanceOrg	newbalanceOrig	nameDest	oldbalanceDest	newbalanceDest	isFraud
0	1	PAYMENT	9839.64	C1231006815	170136.0	160296.36	M1979787155	0.0	0.0	0
1	1	PAYMENT	1864.28	C1666544295	21249.0	19384.72	M2044282225	0.0	0.0	0
2	1	TRANSFER	181.00	C1305486145	181.0	0.00	C553264065	0.0	0.0	1
3	1	CASH_OUT	181.00	C840083671	181.0	0.00	C38997010	21182.0	0.0	1
4	1	PAYMENT	11668.14	C2048537720	41554.0	29885.86	M1230701703	0.0	0.0	0

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 6362620 entries, 0 to 6362619
Data columns (total 10 columns):
 #   Column          Dtype
---  ------          -----
 0   step            int64
 1   type            object
 2   amount          float64
 3   nameOrig        object
 4   oldbalanceOrg   float64
 5   newbalanceOrig  float64
 6   nameDest        object
 7   oldbalanceDest  float64
 8   newbalanceDest  float64
 9   isFraud         int64
dtypes: float64(5), int64(2), object(3)
memory usage: 485.4+ MB
```





VIII. CONCLUSION

The rise of online payment systems has made financial transactions more convenient but has also increased the risk of fraud. This project demonstrates how machine learning techniques can be effectively applied to detect and prevent fraudulent transactions in real time. By leveraging algorithms such as Logistic Regression, Decision Trees, Random Forest, and XGBoost, the proposed system can accurately classify transactions as legitimate or fraudulent, significantly reducing financial losses and enhancing security. The integration of data preprocessing steps, including handling class imbalance and feature scaling, ensures improved model performance. Additionally, the system's ability to continuously

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 5, May 2025

|DOI: 10.15680/IJIRSET.2025.1405503|

learn and adapt to new fraud patterns makes it a robust solution for evolving threats. Overall, this project highlights the importance of intelligent, data-driven approaches in strengthening the trustworthiness of online payment platforms.

IX. FUTURE ENHANCEMENT

The online payment fraud detection system can be enhanced by incorporating deep learning techniques such as recurrent neural networks (RNNs) and convolutional neural networks (CNNs) to better capture complex patterns and sequential transaction behavior. Integrating real-time data streaming platforms like Apache Kafka can improve the system's capability to process and analyze transactions instantly at a larger scale. Additionally, implementing hybrid models that combine supervised and unsupervised learning can help detect previously unseen fraud types more effectively. The system can also be extended to include multi-factor authentication and biometric verification for added security. Finally, expanding the model to support multiple payment methods and international transactions will make the system more versatile and applicable in a global context.

REFERENCES

1.Machine Learning for Cybersecurity Cookbook" by Emmanuel Tsukerman
- A practical guide on applying machine learning techniques for cybersecurity, including fraud detection.

2."Data Mining: Concepts and Techniques" by Jiawei Han, Micheline Kamber, and Jian Pei
- Covers fundamental data mining techniques and their applications, including anomaly decection.

3."Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow" by Aurélien Géron
- Comprehensive guide to machine learning algorithms and implementation using Python tools.

4."Fraud Analytics Using Descriptive, Predictive, and Social Network Techniques" by Bart Baesens
- Focuses specifically on fraud detection methods using data analytics and machine learning.

5."Pattern Recognition and Machine Learning" by Christopher M. Bishop
- Provides the theoretical foundation for machine learning models often used in classification problems.

WEBSITE REFERENCES

1. https://www.geeksforgeeks.org/online-payment-fraud-detection-using-machine-learning-in-python/
2. https://stripe.com/resources/more/how-machine-learning-works-for-payment-fraud-detection-and-prevention
3. https://www.zoho.com/payments/academy/fraud-and-risk-management/machine-
4. https://www.airwallex.com/us/blog/machine-learning-payment-fraud-detection
5. https://www.kaggle.com/datasets/jainilcoder/online-payment-fraud-detection
6. https://github.com/Projects-Developer/ONLINE-PAYMENT-FRAUD-DETECTION-USING-MACHINE-LEING
7. https://medium.com/@sujaychatterjee/anomaly-detection-in-payment-fraud-using-machine-learning-2350bc41003
8. https://towardsdatascience.com/building-a-payment-fraud-detection-model-with-machine-learning-24a4bc6e5f56
9. https://developer.ibm.com/articles/detecting-fraud-in-payment-transactions-using-machine-learning/
10. https://www.datacamp.com/tutorial/payment-fraud-detection-machine-learning

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi cross ref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 6381 907 438 ijirset@gmail.com

www.ijirset.com

Scan to save the contact details