

Online Payment Fraud Detection using Machine Learning Techniques

Machine learning Project

**TEAM NAME:**

Crusaders

**TEAM MEMBERS:**

Panidapu Meghana (23BCE8566) Chiluvuri Sai Sri Vaishnavi (23BCE9684)

**INSTRUCTOR:**

Prof. Sucharitha Jackson

School of Computer Science and Engineering

Online Payment Fraud Detection using Machine Learning Techniques

**Abstract**

When it comes to the simplicity of making a payment while sitting anywhere in the world, online payments have been a source of attractiveness. Over the past few decades, there has been an increase in online payments. E-payments enable businesses earn a lot of money in addition to consumers. However, because elec- tronic payments are so simple, there is also a risk of fraud associated with them. A consumer must ensure that the payment he is paying is going exclusively to the appropriate service provider. Online fraud exposes users to the possibility of their data being compromised, as well as the inconvenience of having to report the fraud, block their payment method, and other things. When businesses are involved, it causes some issues; occasionally, they must issue refunds in order to keep custom- ers. Therefore, it is crucial that both consumers and businesses are aware of these internet scams. A model to determine if an online payment is fraudulent or not is put forth in this study. To determine if a certain Online payment is fraudulent or not, some features like the type of payment, the recipient’s identity, etc. would be taken into account.

# Introduction

Online payments have become more popular during the last few decades. This is because it’s so simple to send money from anywhere, but the pandemic has also contributed signi- ficantly to the rise in e-payments. Numerous studies have demonstrated that e-commerce and online payments will continue to grow in popularity in the years to come. The risk of online payment fraud has also increased as a result of this rise in online payments. Online payment fraud has been shown to have increased over the past few years, making it crucial for consumers and service providers to be aware of these frauds. It is crucial for users to be certain that the payments they make are going to the legitimate recipients; otherwise, they run the risk of having to report fraud, freeze their payment method, and run the chance of having their data shared with criminals, which could occasionally result in more crimes. On the other hand, it’s crucial for companies to check that their cus- tomers aren’t giving money to these fraudsters. Businesses may have to repay money to clients in order to keep their patronage, which puts a strain on them. Even though firms have created and introduced numerous fraud detection programs, only a small number of them are effective in identifying online payment fraud. Although companies make every effort to make the payment method as secure as possible, fraudsters occasionally manage to circumvent security measures and commit these online payment scams. According to

studiesZanin et al. (2018), from 2014 to 2017, the cumulative losses from fraudulent bank card transactions increased globally. Other studiesKalbande [et al.](#_bookmark5) [(2021)](#_bookmark5) concentrate on idea drift, which refers to the possibility of change in the dataset’s underlying distribu- tion over time. Similar to how cardholders or users may alter their purchasing patterns over time, these fraudsters may modify their tactics. These fraudsters are always aware of the customers’ payment methods and behavior, but occasionally their tactics become outdated with time as some professionals work round-the-clock to uncover these scams and shield people from them. Fraud is an illegal technique to obtain somethingYan et al. (2021), thus it’s important to have in place an effective fraud detection system (FDS) that keeps track of all transactions and looks for any indication of fraud. The investig- ator looks into these potentially fraudulent transactions and reports back on whether or not the transaction was actually fraudulent. Machine learning techniques were used to determine if a transaction is fraudulent or not. Data mining techniques were typically used to analyze the patterns of fraudulent and non-fraudulent transactionsWang et al. (2015). Therefore, by analyzing the patterns of the data, a combination of data mining techniques and machine learning can be utilized to determine if transactions are fraudu- lent or real.

Hence, the research question for this study would be- “How far may machine learning methods be utilized to determine whether a specific online transaction is fraudulent or not based on selected features?”.

# Related Work

The synthesis of models for identifying fraudulent transactions in online payments is a challenging task given the large number of studies that have been conducted using various data sets and produced disparate results. These involve determining the issues with obtaining the data, getting the extracted data ready, selecting the appropriate processing procedure, deciphering the outcomes, and analyzing the same [Kolodiziev et al.](#_bookmark7) [(2020).](#_bookmark7) The actual payment data is incredibly sensitive because it contains very sensitive private information on customers or users, and only the companies in charge of handling these data are permitted access to this data. Ranjan et al. (2022)

## Research on the variables that affect credit card theft in online transactions

It is crucial for consumers and organizations to ensure the security of transactions and the private data they include in today’s world, when there are numerous transactions occurring every second. Rai [and Dwivedi](#_bookmark8) [(2020)](#_bookmark8) These thefts are typically carried out when utilizing credit cards to make purchases. This study focuses on credit card fraud detection when making online purchases.

## Protection of privacy

Numerous research are being conducted using the data in a way that protects privacy. One of the experiments was carried out using blockchain technology and machine learning techniques, according to [Kalbande et al.](#_bookmark5) [(2021).](#_bookmark5) The usage of block chain technology, however, can be helpful in protecting the privacy of the data, but we cannot ignore the fact that it is a decentralized solution and has some drawbacks along with it, such as scalability issues and high energy consumption. A supervised machine learning strategy

utilizing block chain technology was developed by Thennakoon et al. (2019). Ethereum was employed by the author to implement block chain technology. 300,000 accounts were used in the study, and the outcomes were compared with a number of machine learning techniques. Additionally, studies utilizing the federated learning and gossip learning paradigms were conducted. [Kolodiziev et al.](#_bookmark7) [(2020)It](#_bookmark7) was determined that gossip learning is unsuccessful because it lacks a central management structure. However, because of its semi-decentralized nature, the federated learning technique, also known as the F.L., was deemed effective and performed better. According to [Jain et al.](#_bookmark4) [(2020),](#_bookmark4) the two transfers where frauds are most prevalent are Cash Out and the ones where money is transferred to a merchant before being transferred to users or occasionally, unknowingly, to fraudsters. The first transfer involves money being transferred from one user to another, a fraudster, or a customer. The second transfer is where frauds are most prevalent. In their assessment of various machine learning algorithms for the identification of frauds when using credit cards, Yee et al. (2018) Accuracy, precision, and specificity criteria were used to assess how well each technique performed in the study. This research also focuses on suggesting a model that makes use of the supervised Random Forest algorithm to improve the precision of identifying credit card payment fraud. Random Forest produced results with a precision of 99.7 percent and an accuracy of 96.2percent. Thennakoon et al. (2019) study of the fraud detection system, which consists of three key parts: the data warehouse, the API module, and the fraud detection models In this study, each of these elements played a simultaneous role. Real-time transactions must be passed between the fraud-detection model, data warehouse, and GUI via the API Module. The output of the machine learning models and the real-time transactions are both stored in a data warehouse. The GUIs, which display fraud alerts for real-time transactions, make it simple for consumers to interact with the fraud detecting system. When the model identifies a specific transaction as fraudulent, a message is displayed on the API module. Additionally, the user receives a message from the API, and their feedback is saved for later examination. The study by Singh et al. (2021) suggests an unique fraud detection system that can identify four different types of fraudulent transactions using the most appropriate algorithm. Rambola et al. (2018) studied the blending of data from many databases and stored it in a suitable format so that data mining techniques could be used on it. The information can be used by any organization where decision-making is necessary in order to identify the fraudulent transactions after further analysis of the data. Neural networks are also used in this study to identify fraudulent transactions.

The sum of all bank transactions is combined to identify clusters. According to Zanin et al. (2018), similar customers are grouped together to have the same data for similar types of customers, which makes it easier to analyze the data further. For example, if one customer or user of a particular area and job requests some specific service from the financial institute. The consumer behavior of each individual user is examined by [Bahnsen](#_bookmark0) [et al.](#_bookmark0) [(2016)](#_bookmark0) as they build a fraud detection model. Additionally, it shows that by pre- processing the data and including recent user behavior, the model’s performance improves by roughly 200 percent when compared to the raw data from the transactions. Wang et al. (2015). This work suggests two ensemble learning approaches, OOB and UOB, by using re sampling and a time-decayed measure in order to address the issue of online class imbalance. The sampling rate appears to be commensurate with the degree of imbalance in the data stream. The data distribution turns out to be a significant element determining the models’ performance. Saputra and Suharjito (2019). Most of these fraud detection systems produce positive results, but many of them produce false positive results, or in other words, they flag particular transactions as fraudulent even though they are not.Be[hera and Panigrahi](#_bookmark1) [(2015)](#_bookmark1) . Companies find it difficult since they don’t want their customers to feel constrained all the time. This study suggests using neural networks and fuzzy clustering. It groups related data sets using the fuzzy c- means clustering technique, and then employs the neural network technique to reduce the classification rate. The investigation is conducted by using the fuzzy c-means technique to input vectors.

## Study on factors influencing frauds in online transaction using credit cards

In order to predict fraudulent transactions, Singh et al. (2021) concentrated on using machine learning techniques like KNN, SVM, and Random Forest. When compared to the other algorithms employed in this study, Random Forest comes out to be the most accurate, with a 99.9 percent accuracy rate. Random forest also comes out to have the lowest rate of false alarms related to fraudulent transactions. Although [Jain et al.](#_bookmark4) [(2020)](#_bookmark4) did not use real-time data, it may still be useful in the future to help organizations like banks become aware of these scams. [Ileberi et al.](#_bookmark3) [(2021)applied](#_bookmark3) the AdaBoost method in addition to some supervised machine learning techniques like logistic regression, decision trees, and support vector machines (SVM) (2021). AdaBoost was primarily used to increase the performance of classifiers when employed singly in terms of performance metrics like accuracy and area under the curve (AUC). The XGB-AdaBoost produces an MCC of 0.99 and addresses the dataset’s class imbalance as well. It might be argued that using AdaBoost has a beneficial impact on machine learning models. ? It focuses on the confusion matrix, which serves as a basis for performance measurement, and studies the identification of fraudulent transactions when using a credit card. Multiple measurements are provided, including sensitivity, specificity, accuracy, and error rate. In this study, Logistic Regression and XGBoost were applied, and it was discovered that the cross validation score for LR was 94.16 percent and the cross validation score for XGBoost was 93.76 percent. While XGBoost’s AUC score was 93.55 percent, LR’s was 94 percent. However, since just 10percent of the transactions in the dataset were used, a lot of information was reported to have been lost in this study, which still has room for development. The outcomes might alter if the same techniques were used on the entire dataset. For financial institutions and IT specialists, resolving the issue of rising

criminality due to online payments is a big challenge. [Kaur et al.](#_bookmark6) [(202](#_bookmark6)1). This study focuses on using machine learning to address this problem. Additionally, it employs a number of data mining algorithms, including CART, C4.5, Naive Bayes, and many others. Using various patterns found by machine learning models, the study seeks to distinguish between legitimate and illegal transactions.Yan et al. (2021)

# Methodology

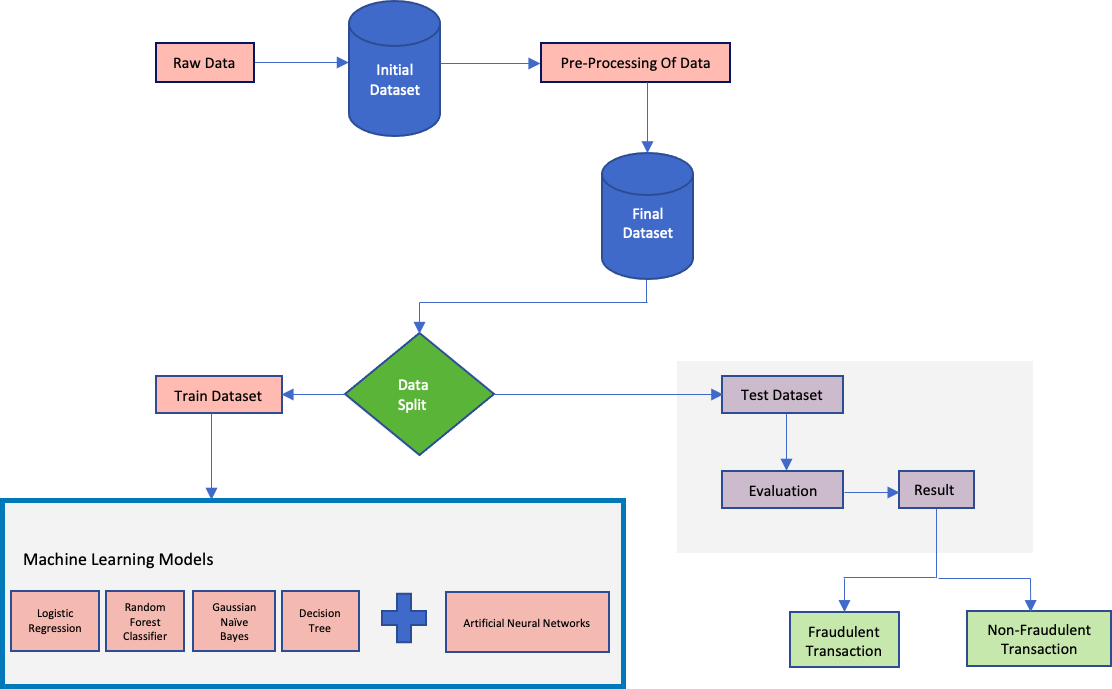
****

Figure 1: Flowchart of Proposed Implementation

When it comes to finding fraudulent online payment transactions, data analysis is crucial. Banks and other financial institutions can adapt the required defences against these frauds with the aid of machine learning techniques. Many businesses and organiz- ations are investing a lot of money in the development of these machine learning systems to determine whether a specific transaction is fraudulent. Machine learning techniques assist these organizations in identifying frauds and preventing their clients who may be at risk for such frauds and occasionally sustain losses as a result.

The research’s data set came from the open platform ”kaggle.” Due to privacy concerns, it is challenging to obtain real-time data sets; therefore, a data collection big enough to conduct the research was taken. The data set has 1048576 records and 11 columns. This data set includes attributes like type (type of payment), amount, ”nameOrig” (customer initiating the transaction), ”pldbalance- Org” (balance before the transaction), ”new- balanceOrig” (balance after the transaction), ”nameDest” (recipient of the transaction),

”pldbalanceDest” (initial recipient balance prior to the transaction), ”newbalanceDest” (the new balance recipient after the transaction), and isFraud which (0 if the transaction is legitimate and 1 if the transaction is fraudu- lent). The figure2. shows all the features

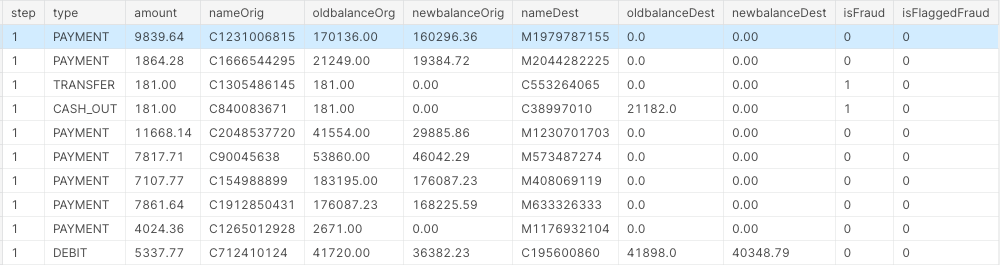


Figure 2: Dataset

in the dataset. Whether a particular transaction is fraudulent or not depends highly on the type of the trasaction, figure 3 below shows the types of transactions and the percentage of the same in our dataset.

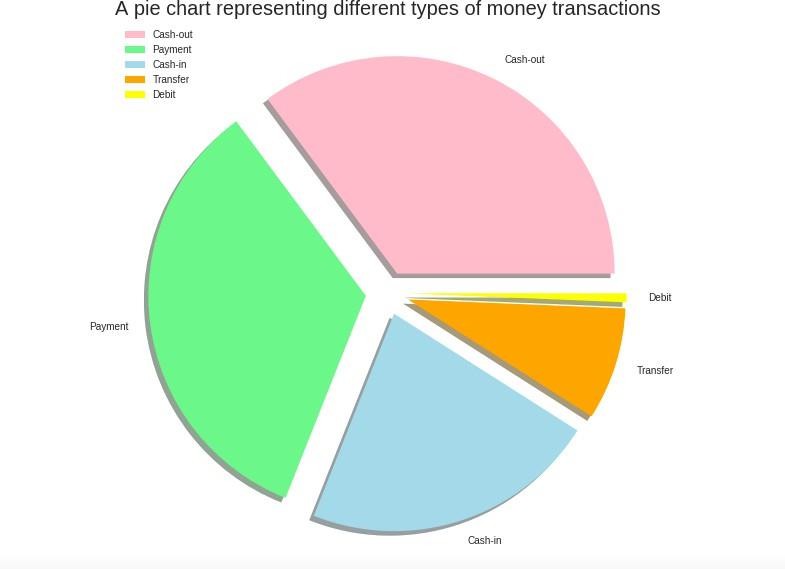


Figure 3: Dataset

## Investigating target distribution:

By looking at the distribution of our target feature ”isfraud” it is clear that there is a huge class imbalance in our data.

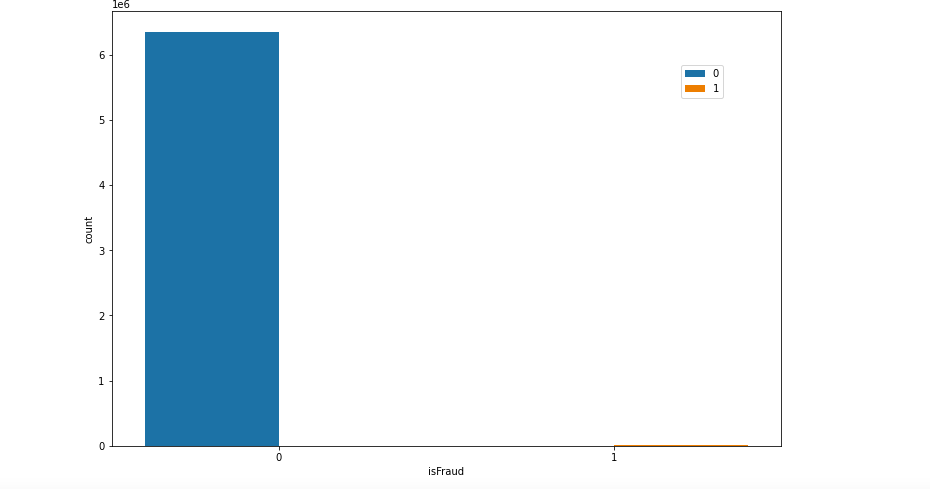


Figure 5: Class Imbalance

As shown in the figure 5 only 8213 transactions are recorded to be fraudulent while 6354407 transactions are recorded to be Non Fraudulent, We would take care of this imbalance between classes in later stages to make sure it does not have any negative impact on our machine learning models.

## Data Preparation:

For the machine learning model to give accurate, high-quality results, the data used to train and test it should be well-prepared. One of the most important steps in data mining is getting the data ready. There are many things to consider, such as how to deal with missing data, duplicate values, removing redundant features from data using correlation matrix and feature selection methods, how to deal with the fact that data isn’t balanced, etc. The quality of the techniques used to prepare the data has a lot to do with how well machine learning works. If the data is not prepared well, it could take a long time to run the models and cost a lot of money. Because of all of these things, the most difficult and time-consuming part of the data mining process is getting the data ready.

### Feature Selection

Feature selection is one of the approaches that helps models perform even better after data cleansing and feature correlation analysis. This method is used to eliminate unnecessary variables, which leads to a smaller feature space and could improve the performance of the model.

In our dataset two features ”namedest” and ”nameOrig” were of less significance as compared to other features, however to compare the same we will be running the models without these features and then including these two features.

## Modelling Approach

Modelling is a very important aspect in machine learning. After the final data pre- paration, which includes steps like handling the class imbalance and feature selection, the proposed models are implemented on the processed or prepared data. The detailed explanation and working of the proposed models are discussed in this section:

### Logistic Regression

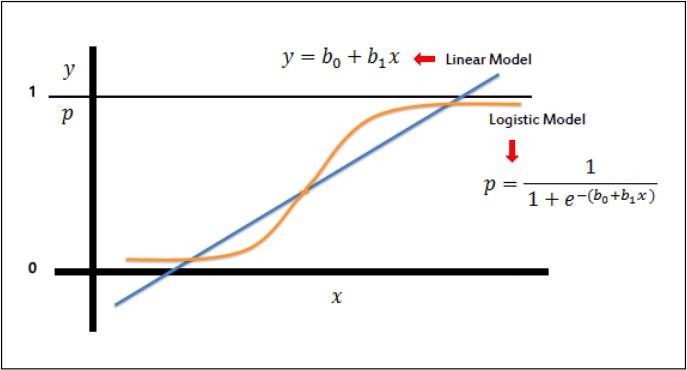
Logistic Regression is the classification of algorithm into multiple categorical values. It includes the use of multiple independent variables which are used to predict a particular outcome of a variable which is dependent on all the independent variables use to train the model. Logistic Regression is similar to linear regression, it predicts a target field rather than a numeric one Zanin et al. (2018). Like predicting True or False, successful or unsuccessful in our case it is fraudulent or non fraudulent. The figure below explains the logistic regression:

Figure 8: Logistic Regression

### Random Forest Classifier

The random forest model is made up of many decision trees that are all put together to solve classification problems. It uses methods like feature randomization and bagging to build each tree. This makes a forest of trees that don’t have anything in common with each other. Every tree in the forest is based on a basic training sample, and the number of trees in the forest has a direct impact on the results.[Bahnsen et al.](#_bookmark0) [(2016)](#_bookmark0)

Tsest fdfdf

### Decision Tree

Decision tree is a supervised machine learning algorithm which uses a combination of rules to make a particular decision, just like a human being. The motive behind decision tree is that one uses the dataset features to create yes or no questions and split the dataset until and unless we isolate all the datapoints those belong to each class.Choi [and](#_bookmark2) [Lee](#_bookmark2) [(2017)](#_bookmark2)

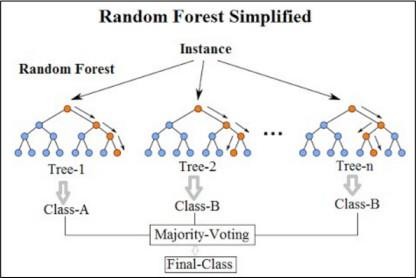


Figure 9: Random Forest

Decision tree is a tree like structure having branch node, leaf node and the root node. The top most node is called the root node.

## 1.2 Artificial Neural Networks

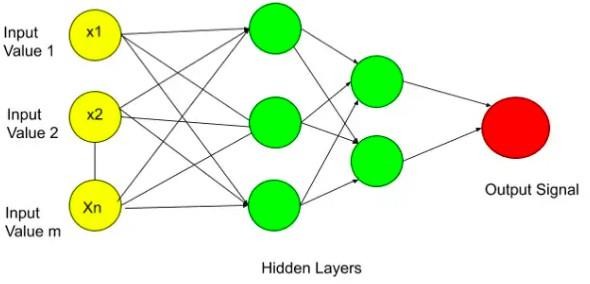
****

Figure 10: Artificial Neural Networks

All the circles in the figure are called neurons and ANN is fully connected with all the neurons.Information is sent to the input layer. The following layer, which is a hidden layer, was then given access to this data by the input layer. The hidden layer carries out certain tasks. then transmit the outcome to the output layer.

# Design Specification

We begin by gathering data from the source, which is followed by pre-processing and EDA (explanatory data analysis) stages. These comprise removing duplicate and null values as well as uncovering hidden patterns in the data. We are filtering our features afterwards to maintain only the columns that are important to our analysis, however for the comparison we are running the models again including all the features which were filtered initially. The training of our baseline models on the training data set came next after we had divided our data into the train, and test datasets.

Figure 11: Confusion Matrix

# Evaluation

The main goal of this research is to use supervised machine learning techniques and Ar- tificial Neural Networks together and see if our proposed method improves the model’s performance more than other state-of-the-art methods.For the comparative study, we did the following two experiments: [1] Figuring out how well a model works by training it on all of the features in the dataset. [2] Figuring out how well the model works by training it on certain features (Eliminating namedest and name orig features).We chose metrics like recall, specificity, F1-score, AUC score, AUC-ROC curve, and the geometric mean of recall and specificity so that we could compare how well our models worked.Because our data isn’t balanced, we can’t judge how well models work based on how accurate they are.Rambola et al. (2018) also compares the results based on the confusion matrix as in this case it is better to compare the True Positive and True negatives and decide based on that that how much accuracy is achieved in our models.The confusion matrix is explained below:

* + True Positive(TP):It shows that the given model has done a good job of figuring out non-fraudulent cases as non-fraud (positive).
  + False Positive(FP): It shows that the model didn’t get the prediction right, fraud- ulent cases were identified as non-fraud (positive).
  + False Negative(FN):It shows that the model didn’t get the prediction right, non- fraudulent cases were identified as fraudulent (negative).
  + True Negative(TN):It shows that the model has been able to accurately predict fraudulent cases as fraudulent (negative).

Precision and specificity show the number of transactions that are considered to be fraud and are frauds. On the other hand, recall/sensitivity values show what percentage of real fraud transactions are correctly classified.F1-score is the average of the notes between Precision and Recall and for better classification, should be close to 1 [Kalbande et al.](#_bookmark5) [(2021).The](#_bookmark5) geometric mean is the sum of both specificity and sensitivity. We chose this metric to judge our models because it works well with unbalanced data [Bahnsen et al.](#_bookmark0)

[(2016).](#_bookmark0) Due to how unbalanced our data is, the most important evaluation metric in our research is the recall and AUC score. We’ve used the confusion matrix to figure out the above evaluation metrics. the equations that follow:- After using the different data

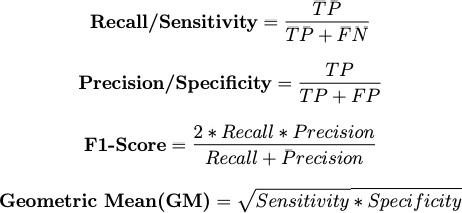


Figure 12:

preparation methods talked about in the above sections, we now have our final dataset to use with our chosen models.To see how well our proposed method of using undersampling for handling Imbalance in data works and eliminating two fetaures(namedest and name- orig),We have done two different case studies for each of the algorithms and used the above- discussed evaluation metrics to rate them.

# Conclusion and Future Work

Online payment fraud has been identified as one of the leading frauds in the past few dec- ades, In this research paper, we discussed the concept of online payment fraud detection. It was seen that feature selection techniques are very important and can be implemented to attain lower false positive rate. We implemented various machine learning algorithms like logistic regression, Random Forest in order to predict if a particular transaction is fraudulent or not. A good fraud detection system should accurately be able to predict if a given transaction is fraudulent or not.

To improve the performance of the models, various techniques such as handling class im- balance, feature selection was used in order to extract the most relevant data. Confusion matrix was used to evaluate the performance of our models, however we did not attain 0 False Positive and false negative score. It is important for a financial organization to attain 0 false positive and negative score as we discussed it impacts on the customer retention and costs lot of money for the refunds. More future works can be done on this research in order to attain the 0 false positive and negative score. Combination of models can be used to attain high accuracy in predicting the transactions as fraudulent and non fraudulent.

# References

Bahnsen, A. C., Aouada, D., Stojanovic, A. and Ottersten, B. (2016). Detecting credit card fraud using periodic features.

Behera, T. K. and Panigrahi, S. (2015). Credit card fraud detection: A hybrid approach using fuzzy clustering neural network.

Choi, D. and Lee, K. (2017). Machine learning based approach to financial fraud detection process in mobile payment system, *IT CoNvergence PRActice (INPRA)* **5**.

Ileberi, E., Sun, Y. and Wang, Z. (2021). Performance evaluation of machine learning methods for credit card fraud detection using smote and adaboost, *IEEE Access* **9**.

Jain, V., Agrawal, M. and Kumar, A. (2020). Performance analysis of machine learning algorithms in credit cards fraud detection.

Kalbande, D., Prabhu, P., Gharat, A. and Rajabally, T. (2021). A fraud detection system using machine learning.

Kaur, P., Sharma, A., Chahal, J. K., Sharma, T. and Sharma, V. K. (2021). Analysis on credit card fraud detection and prevention using data mining and machine learning techniques.

Kolodiziev, O., Mints, A., Sidelov, P., Pleskun, I. and Lozynska, O. (2020). Automatic machine learning algorithms for fraud detection in digital payment systems, *Eastern- European Journal of Enterprise Technologies* **5**.

Rai, A. K. and Dwivedi, R. K. (2020). Fraud detection in credit card data using unsu- pervised machine learning based scheme.

Rambola, R., Varshney, P. and Vishwakarma, P. (2018). Data mining techniques for fraud detection in banking sector.

Ranjan, P., Santhosh, K., Kumar, A. and Kumar, S. (2022). Fraud detection on bank payments using machine learning.

Saputra, A. and Suharjito (2019). Fraud detection using machine learning in e-commerce,

*International Journal of Advanced Computer Science and Applications* **10**.

Singh, P., Chauhan, V., Singh, S., Agarwal, P. and Agrawal, S. (2021). Model for credit card fraud detection using machine learning algorithm.

Thennakoon, A., Bhagyani, C., Premadasa, S., Mihiranga, S. and Kuruwitaarachchi, N. (2019). Real- time credit card fraud detection using machine learning.

Wang, S., Minku, L. L. and Yao, X. (2015). Resampling-based ensemble methods for on- line class imbalance learning, *IEEE Transactions on Knowledge and Data Engineering* **27**.

Yan, T., Li, Y. and He, J. (2021). Comparison of machine learning and neural network models on fraud detection.

Yee, O. S., Sagadevan, S. and Malim, N. H. A. H. (2018). Credit card fraud detection using machine learning as data mining technique, *Journal of Telecommunication, Electronic and Computer Engineering* **10**.

Zanin, M., Romance, M., Moral, S. and Criado, R. (2018). Credit card fraud detection through parenclitic network analysis, *Complexity* **2018**.

4/6/25, 8:49 PM Online Payment Fraud Detection System. - Colab

import pandas as pd import numpy as np

import matplotlib.pyplot as plt import seaborn as sns

from sklearn.model\_selection import train\_test\_split

from sklearn.metrics import classification\_report, confusion\_matrix, accuracy\_score from xgboost import XGBClassifier

from sklearn.preprocessing import LabelEncoder

# Step 1: Generate Synthetic Dataset np.random.seed(42)

n = 5000 # Number of transactions

data = {

'amount': np.random.exponential(scale=200, size=n).round(2),

'location': np.random.choice(['US', 'UK', 'IN', 'CA', 'FR', 'DE'], size=n), 'device': np.random.choice(['Mobile', 'Desktop', 'Tablet'], size=n),

'is\_foreign\_transaction': np.random.choice([0, 1], size=n, p=[0.9, 0.1]),

'is\_high\_risk\_country': np.random.choice([0, 1], size=n, p=[0.95, 0.05]),

'used\_vpn': np.random.choice([0, 1], size=n, p=[0.97, 0.03]), 'failed\_logins': np.random.randint(0, 4, size=n),

'account\_age\_days': np.random.randint(10, 3000, size=n), 'transaction\_hour': np.random.randint(0, 24, size=n),

'is\_fraud': np.random.choice([0, 1], size=n, p=[0.9, 0.1]) # 10% fraud

}

df = pd.DataFrame(data)

# Step 2: Encode Categorical Features le = LabelEncoder()

df['location'] = le.fit\_transform(df['location']) df['device'] = le.fit\_transform(df['device'])

# Step 3: Split into features and target

X = df.drop('is\_fraud', axis=1) y = df['is\_fraud']

# Step 4: Train-test split

X\_train, X\_test, y\_train, y\_test = train\_test\_split( X, y, test\_size=0.3, random\_state=42, stratify=y

)

# Step 5: Train High-Accuracy Model (XGBoost)

model = XGBClassifier(use\_label\_encoder=False, eval\_metric='logloss', random\_state=42) model.fit(X\_train, y\_train)

# Step 6: Predictions

y\_pred = model.predict(X\_test)

# Step 7: Evaluation

print("✅ Accuracy:", accuracy\_score(y\_test, y\_pred))

print("\n📊 Classification Report:\n", classification\_report(y\_test, y\_pred))

# Step 8: Confusion Matrix

sns.heatmap(confusion\_matrix(y\_test, y\_pred), annot=True, fmt='d', cmap='YlGnBu') plt.title("Confusion Matrix")

plt.xlabel("Predicted") plt.ylabel("Actual")

plt.show()

4/6/25, 8:49 PM Online Payment Fraud Detection System. - Colab

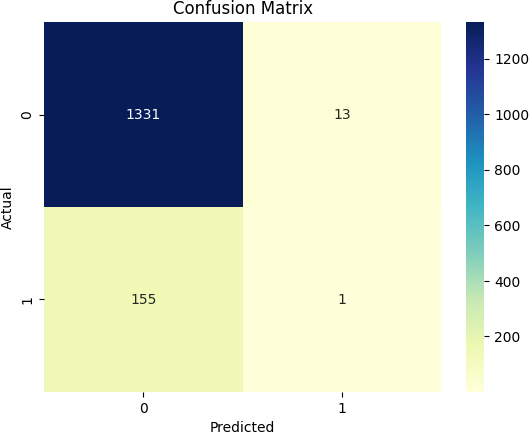
 /usr/local/lib/python3.11/dist-packages/xgboost/core.py:158: UserWarning: [15:17:41] WARNING: /workspace/src/learner.cc:740: Parameters: { "use\_label\_encoder" } are not used.

warnings.warn(smsg, UserWarning)

✅ Accuracy: 0.888

📊 Classification Report:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | precision | recall | f1-score | support |
| 0 | 0.90 | 0.99 | 0.94 | 1344 |
| 1 | 0.07 | 0.01 | 0.01 | 156 |
| accuracy |  |  | 0.89 | 1500 |
| macro avg | 0.48 | 0.50 | 0.48 | 1500 |
| weighted avg | 0.81 | 0.89 | 0.84 | 1500 |



C

C