# ONLINE PAYMENTS FRAUD DETECTION USING MACHINE LEARNING

AN INDUSTRIAL ORIENTED UG PHASE-2 REPORT

Submitted to

**JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY, HYDERABAD**

In partial fulfilment of the requirements for the award of the degree of

**BACHELOR OF TECHNOLOGY**

In

## COMPUTER SCIENCE AND ENGINEERING

Submitted By

**UPPULA DIVYA 19UK1A05F5**

**VEMUNOORI RAMANA 19UK1A05F4**

**DEVA NAGESH 19UK1A05K0**

**VEMURU JAGADEESHWARI 19UK1A05G3**

Under the guidance of

### Mr.G.RAMESH

(Associate Professor)



# DEPARTMENT OF COMPUTER SCIENCE ENGINEERING VAAGDEVI ENGINEEERING COLLEGE

Affiliated to JNTU, HYDERBAD

BOLLIKUNTA, WARANGAL, (T.S)-506005

2019-2023

**DEPARTMENT OF COMPUTER SCIENCE ENGINEERING**

**VAAGDEVI ENGINEEERING COLLEGE**

**BOLLIKUNTA, WARANGAL, (T.S)-506005**



**CERTIFICATE**

This is to certify that the UG Phase-2 entitled “**ONLINE PAYMENTS FRAUD**

**DETECTION USING MACHINE LEARNING** ” is being submitted by **UPPULA**

**DIVYA (19UK1A05F5), VEMUNOORI RAMANA (19UK1A05F4), DEVA NAGESH(19UK1A05K0), VEMURU JAGADEESHWARI(19UK1A05G3)** in partial fulfilment of the requirements for the award of the degree of **Bachelor of Technology** in **Computer Science and Engineering** to **Jawaharlal Nehru Technological University Hyderabad** during the academic year **2019-2023.**

**Project Guide**   **Head of Department**

**Mr. G.RAMESH Dr. R. NAVEEN KUMAR**

(Associate Professor) (Professor)

**External**

# ACKNOWLEDGEMENT

We wish to take this opportunity to express our sincere gratitude and deep sense of respect to our beloved **Dr. P. PRASAD RAO**, Principal, Vaagdevi Engineering College for making us available all the required assistance and for his support and inspiration to carry out this UG Phase-2 in the institute.

We extend our heartfelt thanks to **Dr. R. NAVEEN KUMAR**, Head of the Department of CSE, Vaagdevi Engineering College for providing us necessary infrastructure and thereby giving us freedom to carry out the UG Phase-2.

We express heartfelt thanks to Smart Bridge Educational Services Private Limited, for their constant supervision as well as for providing necessary information regarding the UG Phase-2 and for their support in completing the UG Phase-2.

We express heartfelt thanks to the guide, **Mr. G. RAMESH**, Associate Professor, Department of CSE for his constant support and giving necessary guidance for completion of this UG Phase-2.

Finally, we express our sincere thanks and gratitude to our family members, friends for their encouragement and outpouring their knowledge and experiencing throughout this.

**UPPULA DIVYA 19UK1A05F5**

**VEMUNOORI RAMANA 19UK1A05F4**

**DEVA NAGESH 19UK1A05K0**

**VEMURU JAGADEESHWARI 19UK1A05G3**

## ABSTRACT

In today's world, people depend on online payments for almost everything. Online transactions have their own merits like easy to use, feasibility, faster payments etc., but these kinds of transactions also have some demerits like fraud transactions, phishing, data loss, etc. With increase in online transactions, there is a constant threat for frauds and misleading transactions which can breach an individual's privacy. Hence, many commercial banks and insurance companies devoted millions of rupees to build a transaction detection system to prevent high risk transactions. We presented a machine learning - based transaction fraud detection model with some feature engineering. The algorithm can get experience; improve its stability and performance by processing as much as data possible. These algorithms can be used in the project that is online fraud transaction detection. In these, the dataset of certain transactions which is done online is taken. Then with the help of machine learning algorithms, we can find the unique data pattern or uncommon data patterns which will be useful to detect any fraud transactions. For the best results, the XGBoost algorithm will be used which is a cluster of decision trees. This algorithm is recently dominating this ML world. This algorithm has features like more accuracy and speed when compared to other ML algorithms.

Keywords – Fraud detection, Machine learning, Xgboost algorithm, classification, Data preprocessing, Prediction.

TABLE OF CONTENTS

1. **INTRODUCTION…………………………………………………..1**

# CODE SNIPPETS ………………………………………………2-18

2.1 Model code…………………………………………………...2-14

2.2 Html code and python code…………………………………15-18

# CONCLUSION ………………………………………………...19-21

1. **APPLICATIONS…………………………………………………..22**
2. **ADVANTAGES…………………………………………………….23**
3. **FUTURE SCOPE………………………………………………….24**
4. **BIBILOGRAPHY……………………………………………...27-28**

# 1. INTRODUCTION

In today’s world, we are on the way to become a cashless world. According to various surveys and researches, people performing the online transactions is increased a lot, it’s expected that in future years this will go on increasing. Now, while this might be exciting news, on the other-side fraudulent transactions are on the rise as well. Even due to various security systems being implemented, we still have a very high amount of money being lost due to fraudulent transactions. Online Fraud Transaction can be defined as a case where a person uses someone else’s credit card for personal reasons or for knowing a persons personal info, while the owner and the card issuing authorities are unaware of the fact that the card is being used. Fraud detection involves monitoring the activities of users to estimate, perceive or avoid objectionable behavior, which consists of fraud, intrusion, and defaulting.

The online payment systems has helped a lot in the ease of payments. But, at the same time, it increased in payment frauds. Online payment frauds can happen with anyone using any payment system, especially while making payments using a credit card / debit card. That is why detecting online payment fraud is very important for credit card companies to ensure that the customers are not getting charged for the products and services they never paid.

Most of the E-commerce sites runs on online payments the fraudsters are ready to get the information / personal data once if the fraudster is known the card CVV number or payment UPI-ID then the fraudsters are entering and knowing the personal data of an individual, Even if they know the card number they can predict

CVV number. Because there are many ways now-a-days to predict and various algorithms to predict this may leads to the losing the personal data of a individual without is concern

# 2. CODE SNIPPETS

**2.1 MODEL CODE**



**Figure 1:** .ipynb code importing libraries & mounting dataset from Drive.

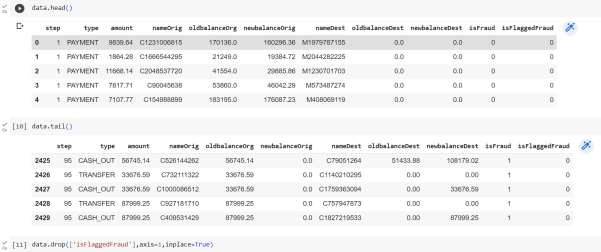
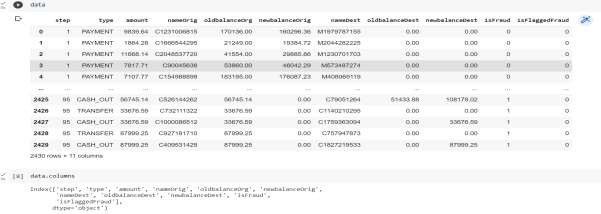


Figure 2: .ipynb code displaying few rows, columns & column names from the dataset.

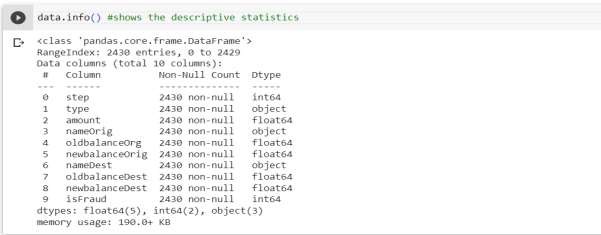


Figure 3: .ipynb code describe in detail info using info() method.



Figure 4: .ipynb code for heatmap shows 2 dimensional representation of dataset.

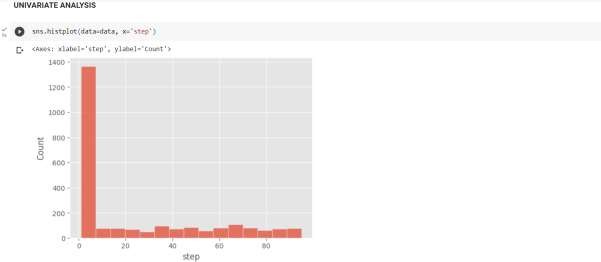


Figure 5: .ipynb code for univariate analysis of step column.

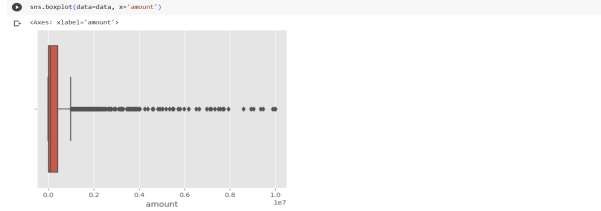
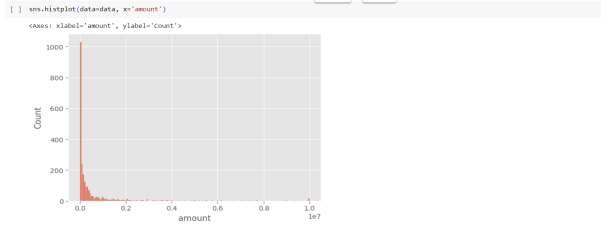
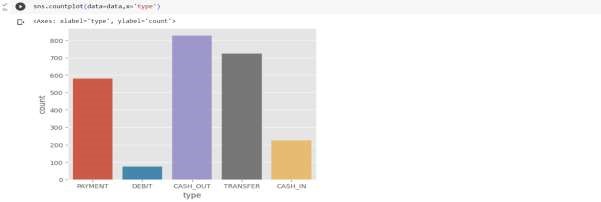
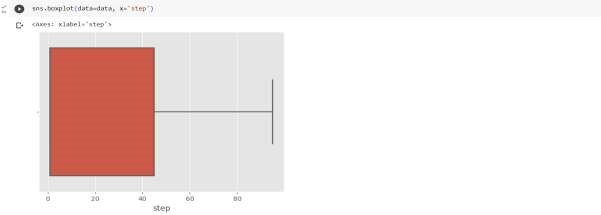


Figure 6: .ipynb code for different columns present in dataset.

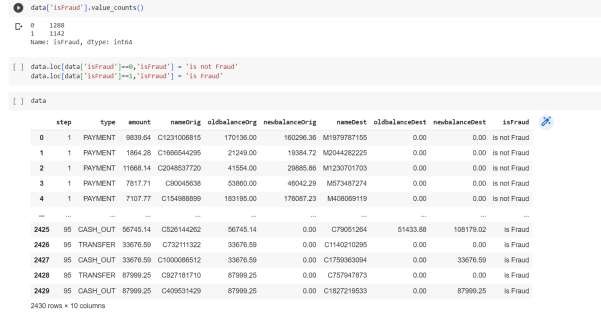


Figure 7: .ipynb code for count of fraud and non fraud transactions & Assigining is fraud=1 & is not fraud=0, displaying dataset.

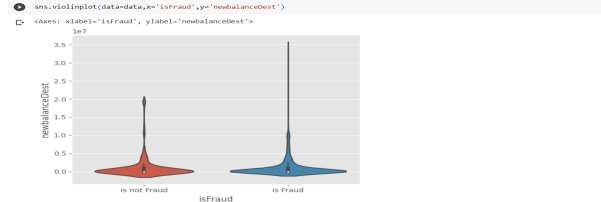
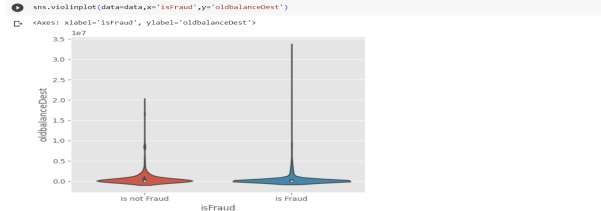
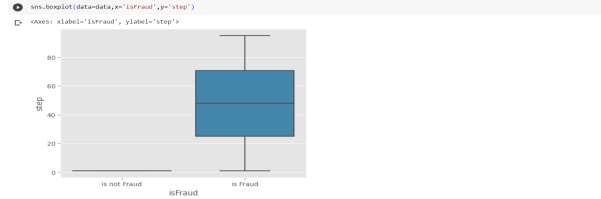
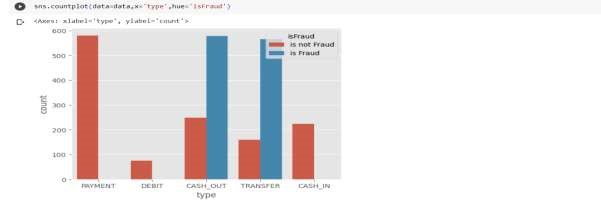
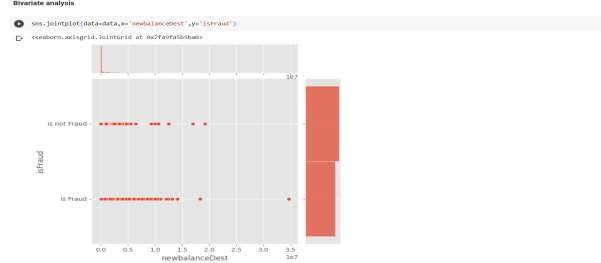


Figure 8: .ipynb code displaying Bi-variate analyasis gives relationship between each variable in dataset.

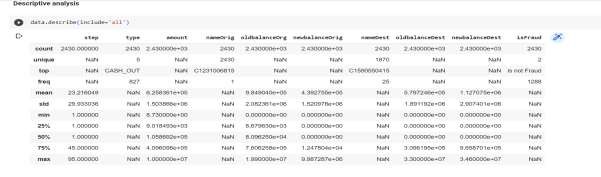


Figure 9: .ipynb code for descriptive analysis it describes the data.

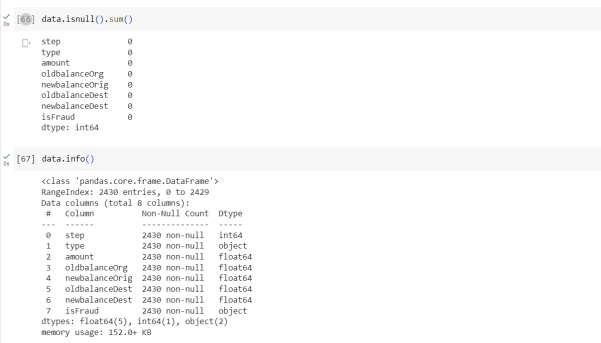
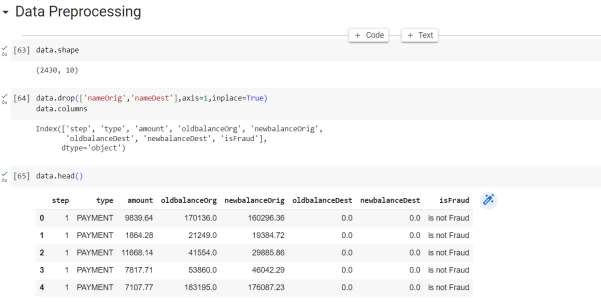


Figure 10: .ipynb code for Data preprocessing, Raw data to processing procedure.



Figure 11: .ipynb code for removing outliers & transformation plot values.

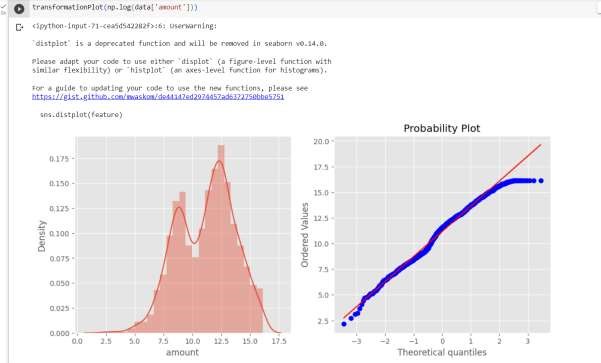


Figure 12: .ipynb code for transformation plot & graphs.

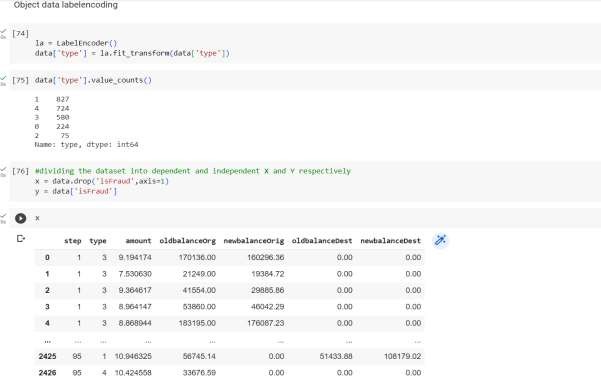


Figure 13: .ipynb code for object label encoding converts categorical values to numerical.



Figure 14: .ipynb code splitting data into train and test.



Figure 15: .ipynb code for Random Forest model.

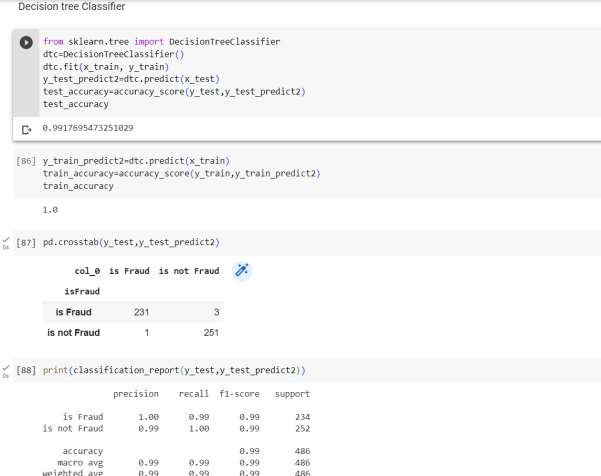


Figure 16: .ipynb code for Decesion tree classifier.

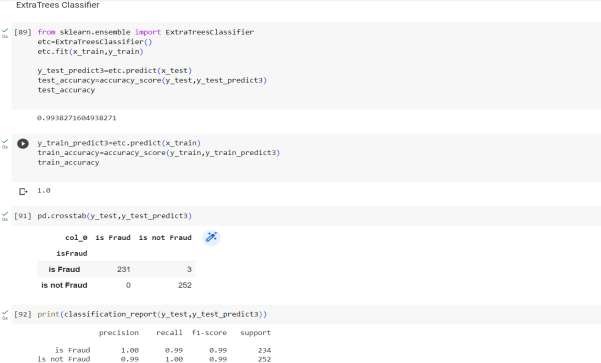


Figure 17: .ipynb code for extra trees classifier.

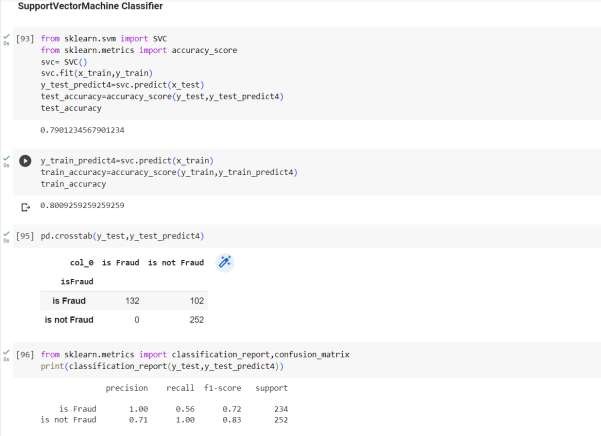


Figure 18: .ipynb code for support vector machine classifier.



Figure 19: .ipynb code for Label encoding converts categorical columns to numerical columns.

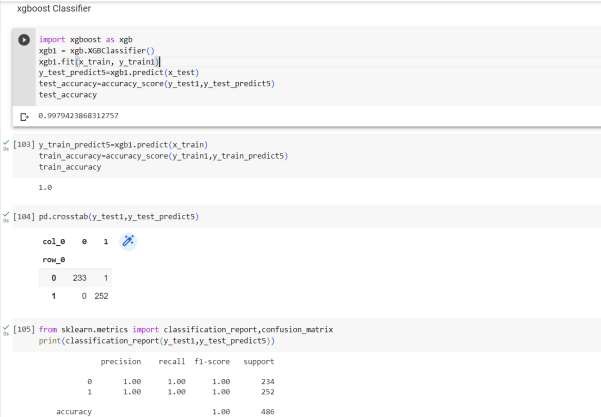


Figure 20: .ipynb code for xgboost classifier.



Figure 21: .ipynb code for comparing the models & accuracy of each model, importing pickle file(.py code).

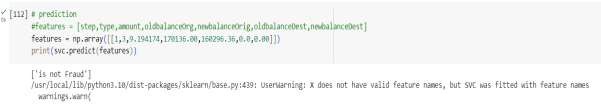


Figure 22: .ipynb code for prediction & predicting by giving values.

**2.2 HTML CODE AND PYTHON CODE**

**1. app.py code:**

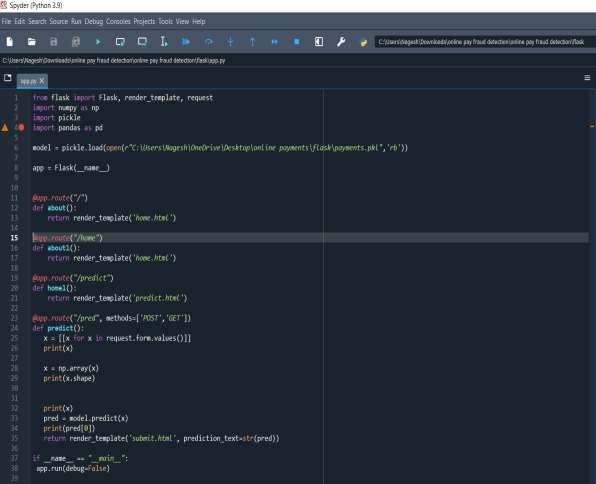


Figure 23: .python code used for rendering all the HTML pages.

**home.html**

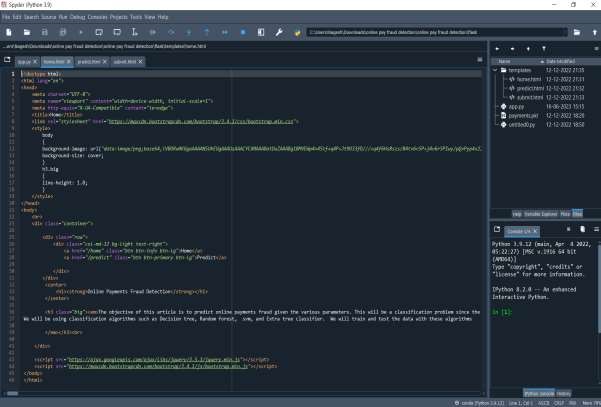


Figure 23: home.html page is the code for homepage of our web application.

**predict.html:**

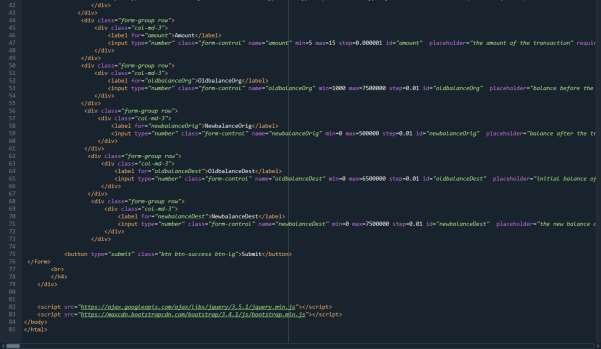
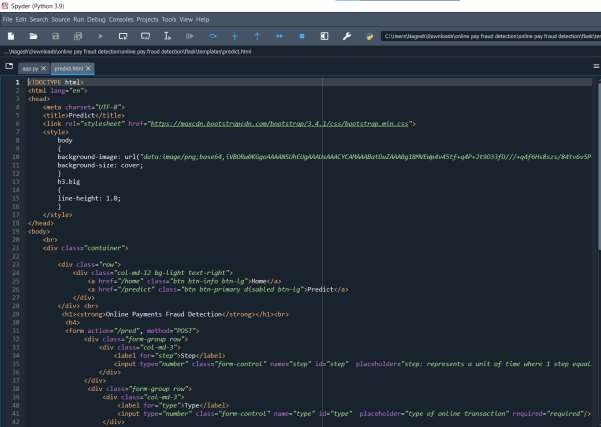


Figure 24: predict.html page which predicts the output. By taking the inputs from user.

**Submit.html**

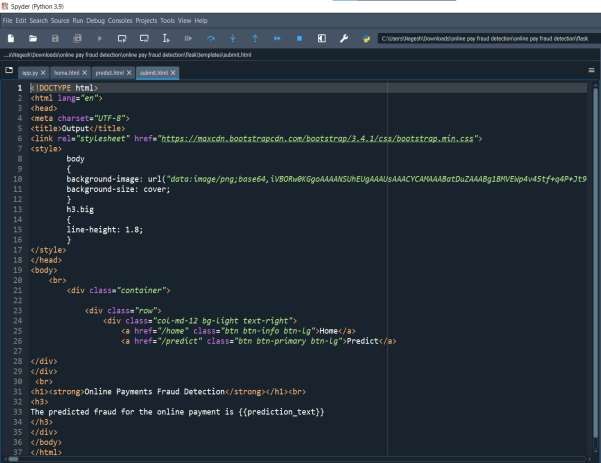


Figure 25: submit.html is a button when we enter values & click on submit button it displays a message associated with code.

# 3. CONCLUSION



**Figure 26:** **Home page (which gives introduction to Online payments Fraud Detection)**



## Figure 27: Input page (which takes input from user)



## Figure 28: Output page (Displays that the payment is fraud)



## Figure 29: Input page (which takes input from user)



**Figure 30: Output page (Displays that the payment is not fraud)**

# 4. APPLICATIONS

The areas where this solution can be applied:

* Bank Transfers & Banking Applications.
* QR codes/UPI payments.
* Digital wallets like phone pe, paytm etc..,
* Swipping machines (card cvv).

# 5. ADVANTAGES

1. **Improved Security:** Online payment fraud detection projects employ advanced algorithms and techniques to identify and prevent fraudulent activities. This helps in enhancing the overall security of online transactions and protects both businesses and customers.
2. **Real-Time Detection:** Online payment fraud detection systems can analyze

transactions in real time, enabling the identification of suspicious patterns or behaviors instantly. This allows for immediate action to be taken, such as blocking a transaction or flagging it for manual review.

1. **Cost Savings:** By implementing an effective fraud detection system, businesses can minimize financial losses due to fraudulent activities. Identifying and preventing fraudulent transactions early on can save significant amounts of money that would otherwise be lost.
2. **Enhanced Customer Trust:** A robust fraud detection system reassures customers that their financial information is secure when making online payments. This helps to build trust and confidence in the business, leading to increased customer satisfaction and loyalty.
3. **Scalability:** Online payment fraud detection systems can handle large volumes of transactions, making them scalable for businesses of different sizes. As the volume of online transactions increases, the system can adapt and accommodate the growing demands.

# 6. DISADVANTAGES

1. **False Positives:** One of the challenges in online payment fraud detection is the occurrence of false positives, where legitimate transactions are incorrectly flagged as fraudulent. This can inconvenience customers and lead to a loss of business if genuine transactions are blocked or delayed.

1. **Evolving Fraud Techniques:** Fraudsters are continually adapting their techniques to bypass detection systems. Keeping up with new and emerging fraud patterns and updating the fraud detection algorithms accordingly can be challenging.

1. **Privacy Concerns:** Online payment fraud detection projects involve the analysis of large amounts of personal and financial data. Ensuring the privacy and security of this sensitive information is crucial to prevent unauthorized access or data breaches.

# 7. FUTURE SCOPE

On our Dataset, we have applied Random Forest, Decision Tree, Xgboost Classifier, SVM, and Extra tree classifier, Xgboost has got the highest accuracy.

**Enhancements that can be made in the future:**

Online payment Fraud Transaction Detection System is basically an extension of the existing system. Using This system, the algorithms which we used to train the dataset and provide the appropriate output. In the long run, this system will be quite beneficial as it provides an efficient system to create a secure transaction system to analyse and detect fraudulent transactions. The Xgboost algorithm is a popular and efficient open-source implementation of the gradient boosted trees algorithm. Gradient boosting is a supervised learning algorithm, which attempts to accurately predict a target variable by combining the estimates of a set of simpler, weaker models. This accuracy can be increased further by providing a huge dataset for model training. The scope of this application is very far reaching. This system can be used to detect the features of fraud transactions in a dataset which is very well applicable in various sectors like banking, insurance, e-commerce, money transfer, bill payments, etc. This will indeed help to increase security.

# 8. BIBILOGRAPHY

1. K.Chaudhary, J.Yadav, “A review of fraud: A comparative study.”decis. Support syst, vol 50, no3, pp.602-613,2011.
2. Katherine J. Barker , Jackie D’Amato ,Paul Sheridon,2008 “Credit card fraud :awareness and prevention”, Journal+- of financial Crime ,Vol. 15issue:4,pp.398-410.
3. “CreditCard Fraud Detection Based on Transaction Be haviour -by John Richard D. Kho, Larry A. Vea” published by Proc. of the 2017 IEEE Region 10 Conference (TENCON), Malaysia, November 5- 8, 2017.
4. Customer Transaction Fraud Detection Using Xgboost Model -by Yixuan Zhang, Ziyi Wang, Jialiang Tong, Fengqiang Gao June, 2020.
5. Wang, M., Yu, J., & Ji, Z. (2018). Credit Fraud Risk Detection Based on XGBoost-LR Hybrid Model.
6. Mishra, C. Ghorpade, “Credit Card Fraud Detection on the Skewed Data Using Various Classification and Ensemble Techniques” 2018 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS) pp. 1- 5. IEEE.
7. [https://thecleverprogrammer.com/2022/02/22/online-payments-fraud-detection-withmachine-learning/](https://thecleverprogrammer.com/2022/02/22/online-payments-fraud-detection-with-machine-learning/)
8. [https://www.geeksforgeeks.org/online-payment-fraud-detection-using-machinelearning-in-python/](https://www.geeksforgeeks.org/online-payment-fraud-detection-using-machine-learning-in-python/)

# 9.HELP LINE

**PROJECT EXCEUTION:**

STEP-1: Go to Google, search google colaboratory & launch.

STEP-2: After launching of collab. STEP-3: Open “Major project .ipynb file.” STEP-4: Then run all the cells.

STEP-5: All the data preprocessing, training and testing, model building, accuracy of the model can be showcased.

STEP-6: And a pickle file will be generated.

STEP-7: Create a Folder named FLASK on the DESKTOP. Extract the pickle file into this Flask Folder.

STEP-8: Extract all the html files (home.html, predict.html, submit.html) and python file(app.py) into the FLASK Folder.

STEP-9: Then go back to ANACONDA NAVIGATOR and the launch the SPYDER.

STEP-10: After launching Spyder, give the path of FLASK FOLDER which you have created on the DESKTOP.

STEP-11: Open the app.py and html files present in the Flask Folder.

STEP-12: After running of the app.py, open ANACONDA PROMPT and follow the below steps: cd File Path< > click enter python app.py< >click enter (We could see running of files).

STEP-13: Then open BROWSER, at the URL area type >> localhost:5000.

STEP-14: Home page of the project will be displayed.

STEP-15: Click on ― Predict. Give the inputs then it will be predict fraud payment or not.