# ONLINE PAYMENTS FRAUD DETECTION USING MACHINE LEARNING

A UG PROJECT PHASE-2 REPORT

Submitted to

## JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY,

**HYDERABAD**

In partial fulfillment 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 esteemed guidance of

**Mr. G. RAMESH**

**(**Associative Professor)



## DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING VAAGDEVI ENGINEERING COLLEGE

(Affiliated to JNTUH, Hyderabad)

Bollikunta, Warangal – 506005 **2019–2023**

# DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING VAAGDEVI ENGINEERING COLLEGE

**BOLLIKUNTA, WARANGAL – 506005**

**2019 – 2023**



**CERTIFICATE OF COMPLETION UG PROJECT PHASE-2**

This is to certify that the UG Project 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

fulfillment 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 **2022-23**, is a record of work carried out by them under the guidance and supervision.

|  |  |
| --- | --- |
| **Project Guide** | **Head of the 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 Project 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 Project 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 Project Phase-2 and for their support in completing the UG Project Phase-2.

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

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

**U. DIVYA 19UK1A05F5**

**V.RAMANA 19UK1A05F4**

**D.NAGESH 19UK1A05K0**

**V.JAGADEESHWARI 19UK1A05G3**

**TABLE OF CONTENTS:**

**1. INTRODUCTION .................................................................... 6**

## 2. CODE SNIPPETS ............................................................... 7-25

**2.1 LOADING THE DATA SET ................................... 7-9**

**2.2 DATA PREPROCESSING ................................... 9-17**

**2.3 MODEL BUILDING............................................ 17-22**

**2.4 APPLICATION BUILDING .............................. 23-25**

1. **CONCLUSION .................................................................... 26**
2. **APPLICATIONS ................................................................. 27**
3. **ADVANTAGES ................................................................... 27**
4. **DISADVANTAGES ............................................................ 27**
5. **FUTURE SCOPE ................................................................ 28**
6. **BIBLIOGRAPHY ............................................................... 29**
7. **HELP FILE ......................................................................... 30**

# LIST OF FIGURES PAGE NO

**Figure 1:** Importing important libraries .........................................................................7

**Figure 2:** Reading the dataset ....................................................................................... 8

**Figure 3:** Descriptive analysis of dataset ...................................................................... 9

**Figure 4:** Checking the dataset ................................................................................... 10

**Figure 5:** Checking dataset for null values ................................................................. 11

**Figure 6:** Merging the dataset columns ...................................................................... 11

**Figure 7:** Handling negative values .............................................................................12

**Figure 8:** Exploratory data analysis .............................................................................13

**Figure 9:** Correlation matrix of dataset ...................................................................... 14

**Figure 10:** Handling categorical values ......................................................................16

**Figure 11:** Splitting data into train and test ................................................................ 17

**Figure 12:** Random forest regressor ........................................................................... 17

**Figure 13:** Decision tree regressor .............................................................................. 18

**Figure 14:** XgBoost model… ..................................................................................... 19

**Figure 15:** ARIMA model… ...................................................................................... 21

**Figure 16:** Comparing the models .............................................................................. 22

**Figure 17:** HTML code................................................................................................ 23

**Figure 18:** Python code ............................................................................................... 24

**Figure 19:** Web application link generation ............................................................... 24

**Figure 20:** Web UI Page ............................................................................................. 25

**Figure 21:** Inputs to Web application ......................................................................... 26

**Figure 22:** Predicted sales units ..................................................................................26

# 1. INTRODUCTION

In today’s world, we are on the verge to become a cashless world. According to various surveys and researches, people performing online transactions has 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 while the owner and the cardissuing authorities are unaware of the fact that the card is being used. Fraud detection involves monitoring the activities of populations of users to estimate, perceive or avoid objectionable behavior, which consists of fraud, intrusion, and defaulting. Most of the time, a person who has become a victim of such fraud doesn't have any idea about it until the very end.

UG Project Phase-2 involves all the coding and implementation of the design which we have retrieved from UG Project Phase-1. All the implementation is done and conclusions are retrieved in this phase. We will also work on the applications, advantages, and disadvantages of the project in this phase. Future scope of the project will be also discussed in the UG Project Phase-2.

# 2. CODE SNIPPETS

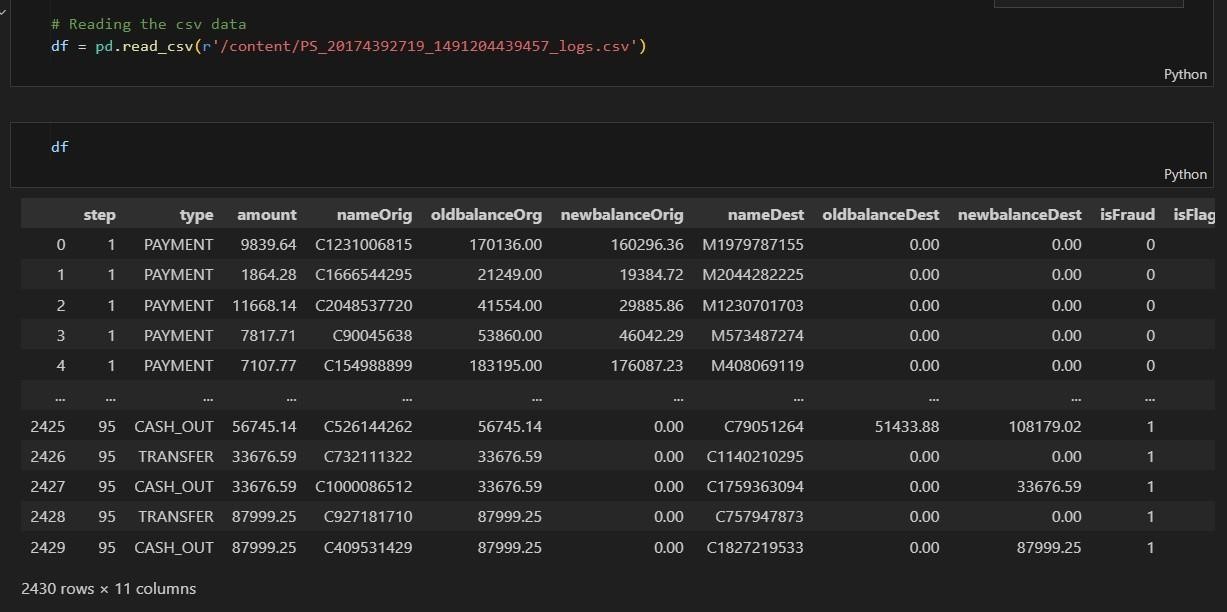
**2.1 LOADING THE DATASET:**

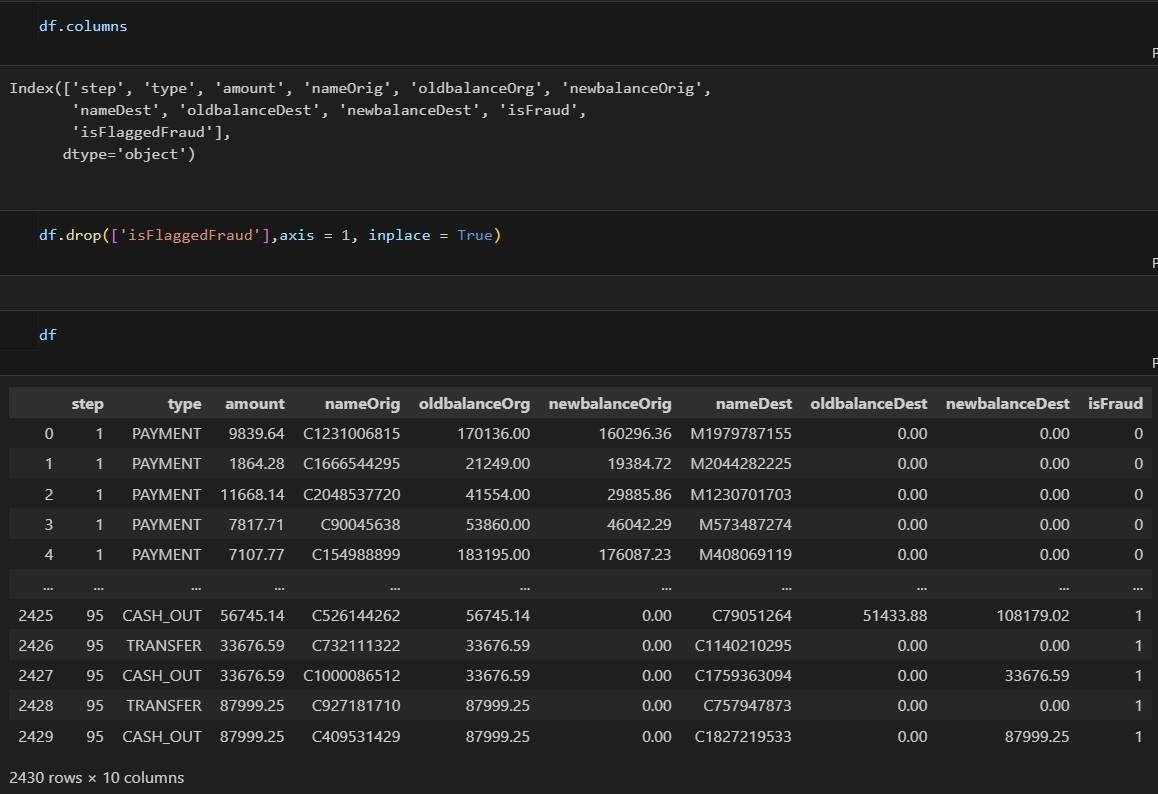
## Activity 1: Importing Libraries



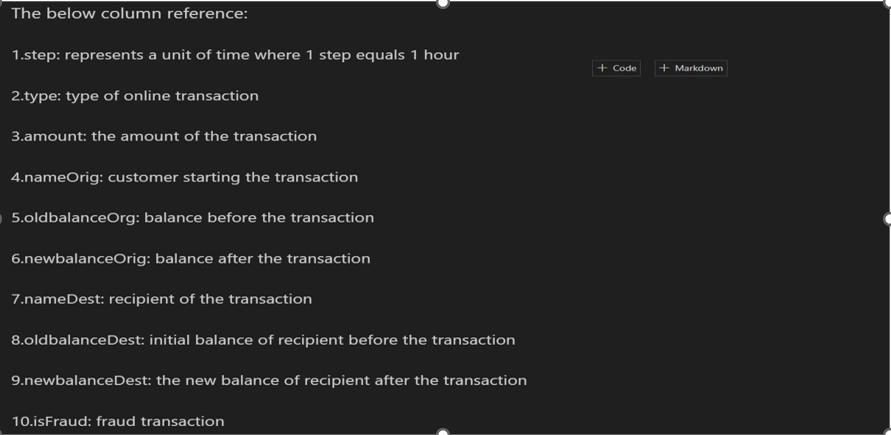
**Figure 1 : importing required libraries**

### Activity 2 : Reading the csv data

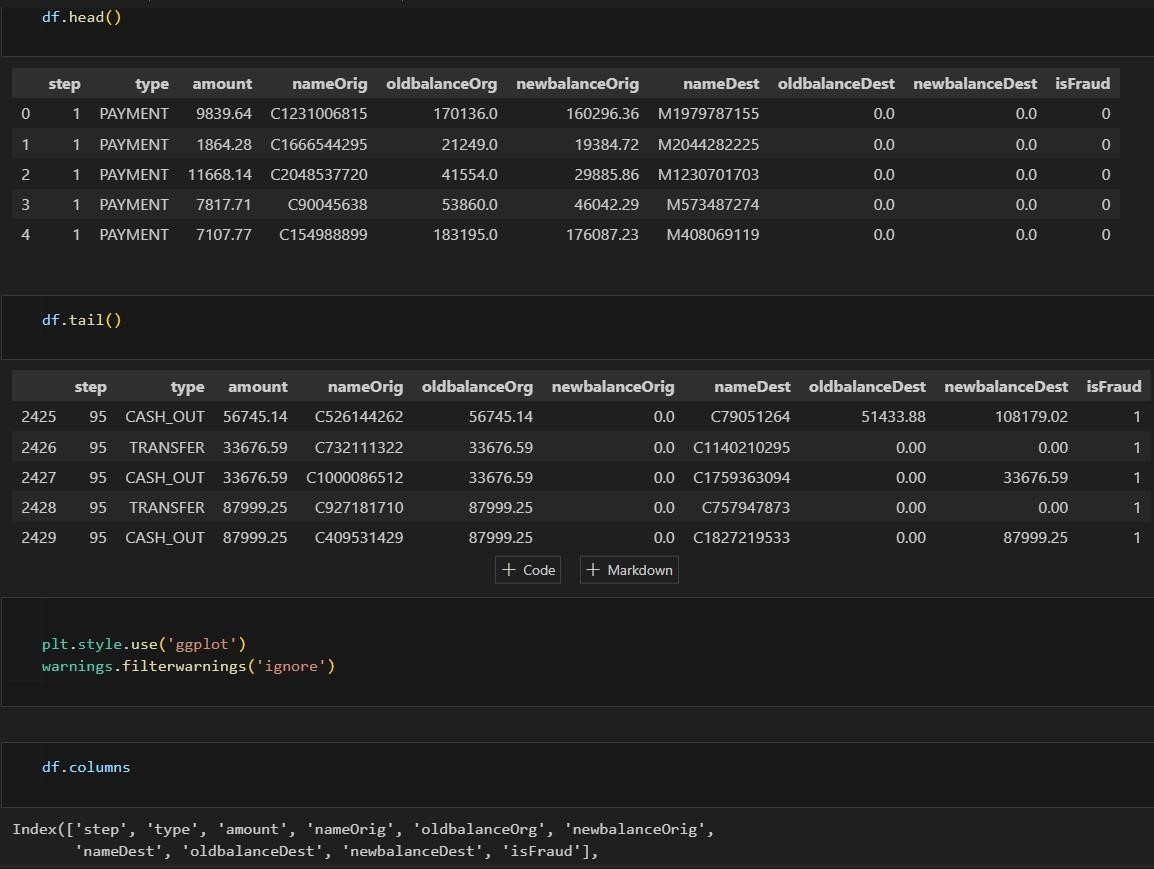




**Figure 2 : Reading the csv data About the data set**



**Figure 3 : About the dataset**



### Checking for Co relation

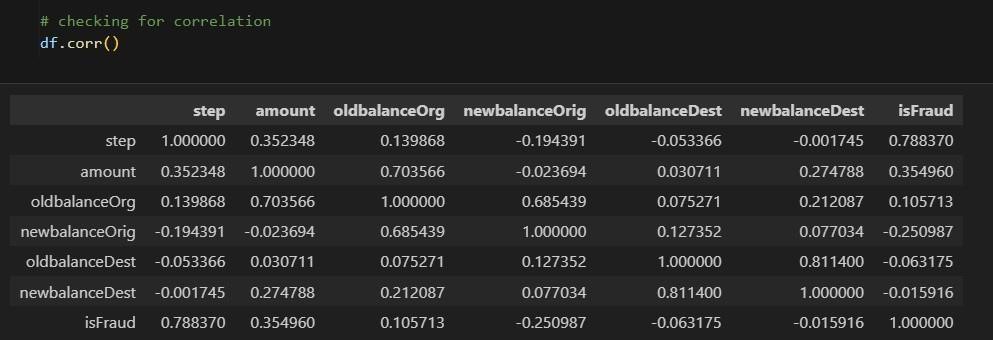
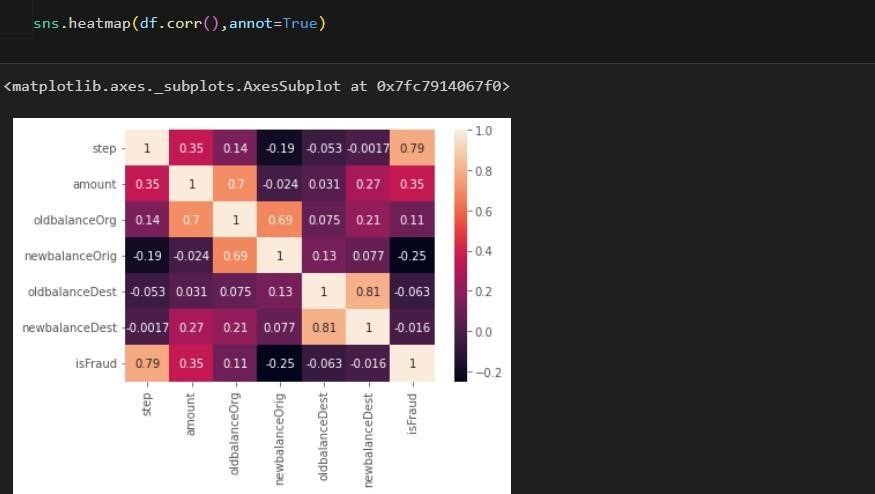
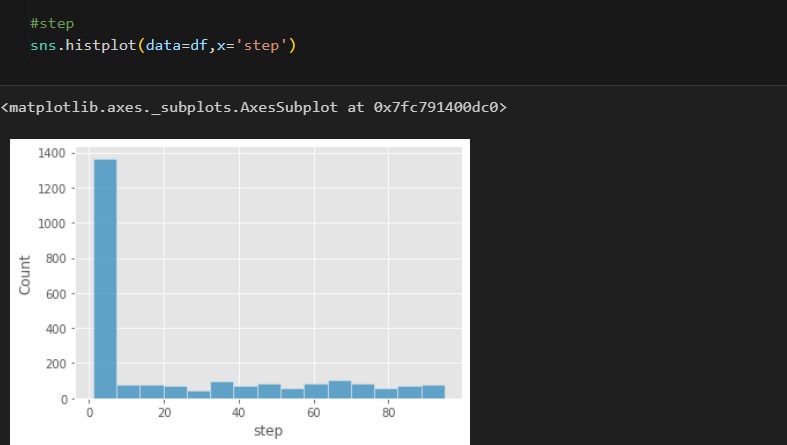


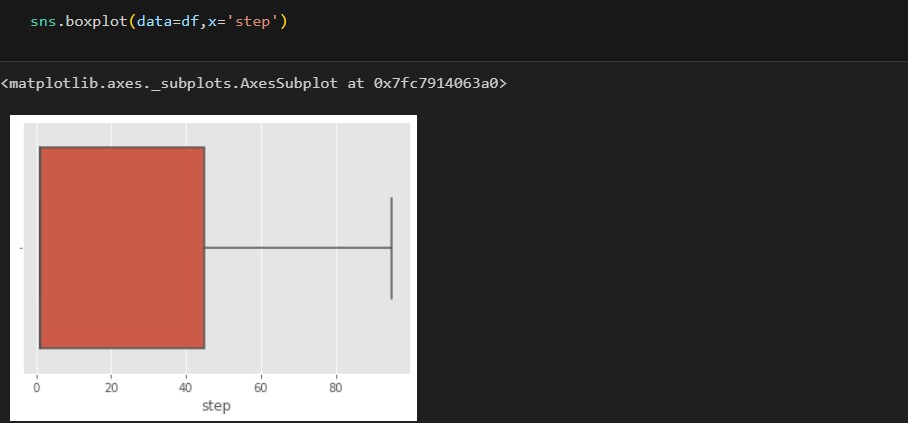
Figure 4 :checking correlation of dataset using d

### Heat map

S

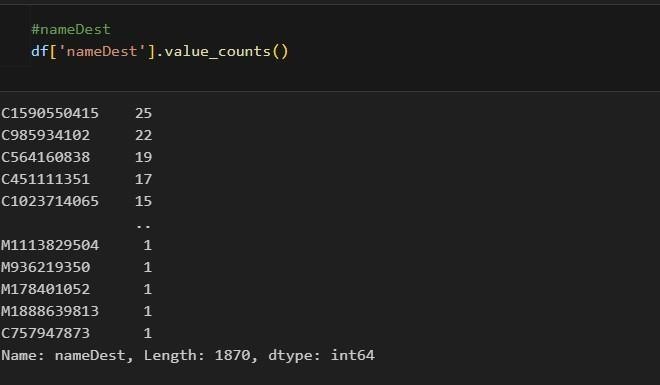
**Activity 3 : Univariate Analysis**





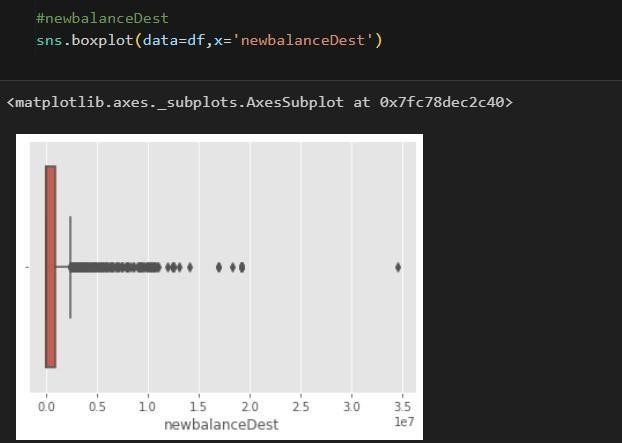
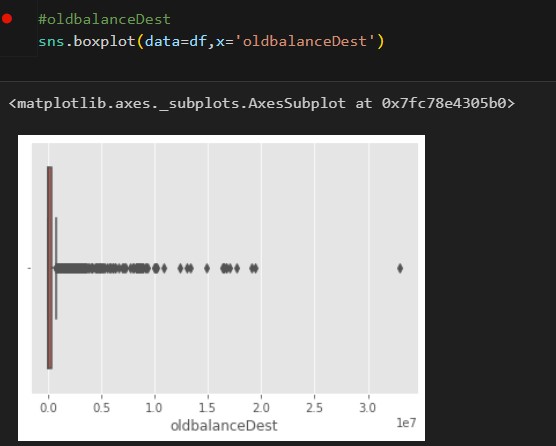




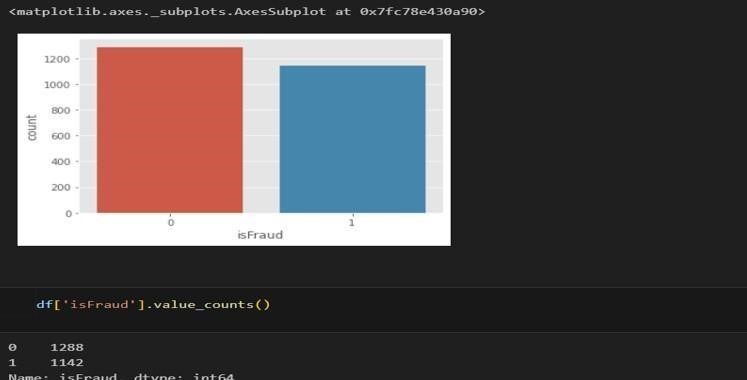


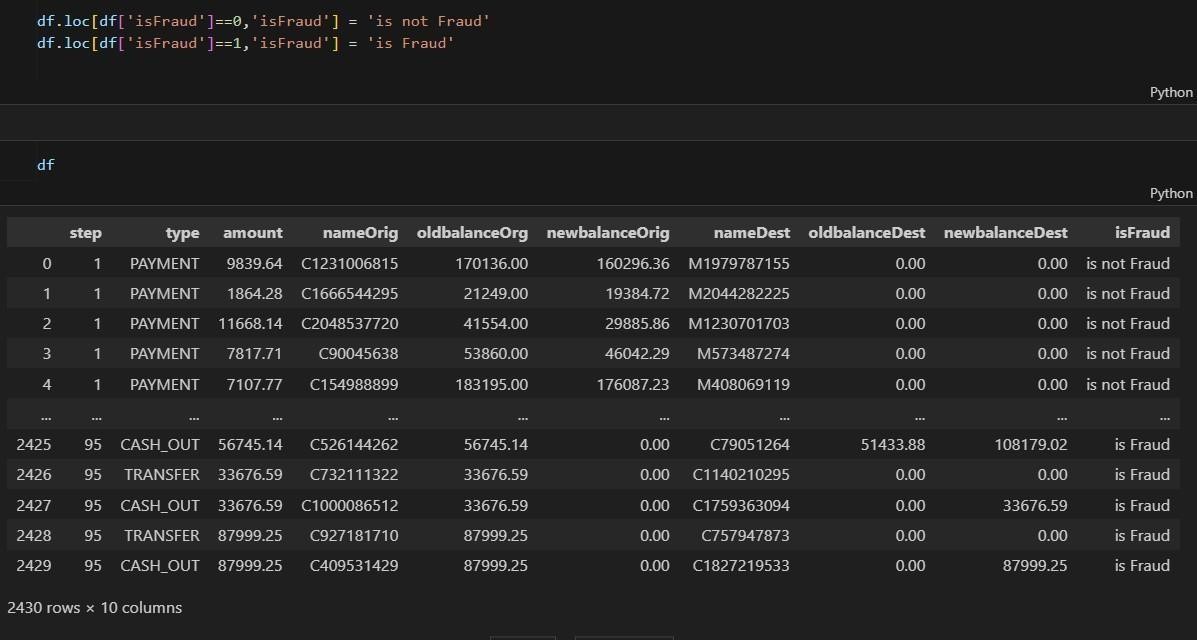
n

FIGURE

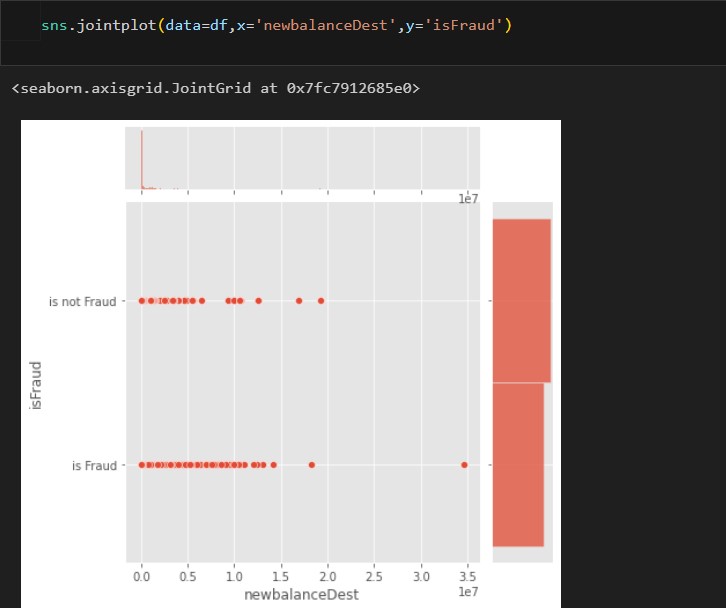


FIGURE





**ACTIVITY 4 :**





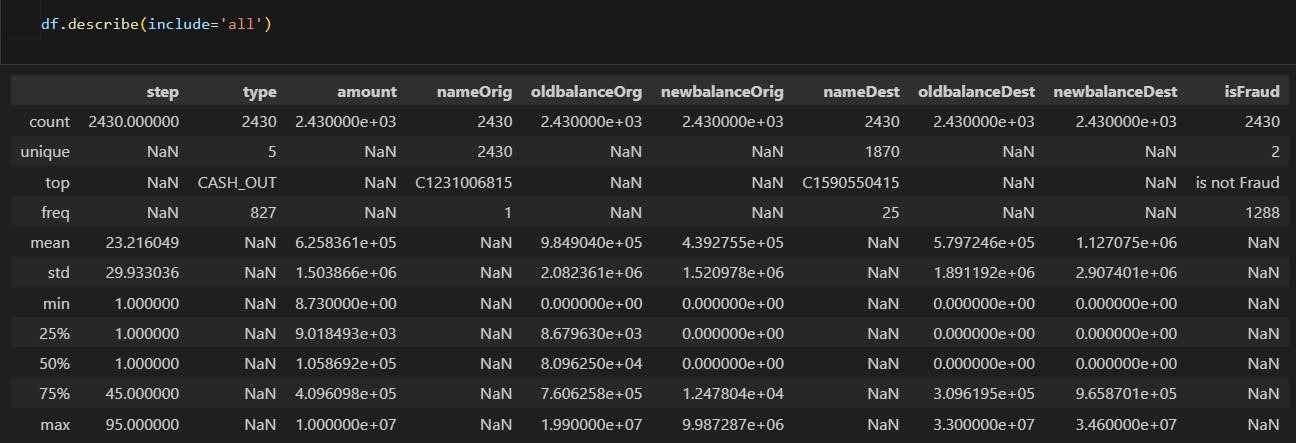
FIGURE



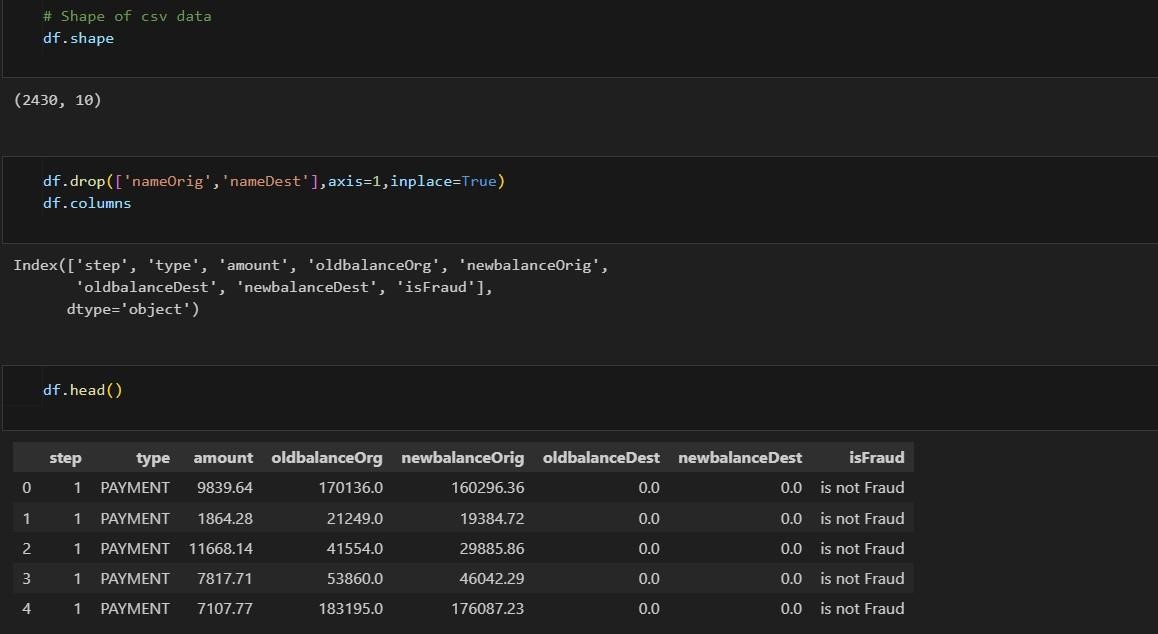
FIGURE



### Activity 5: Descriptive analysis

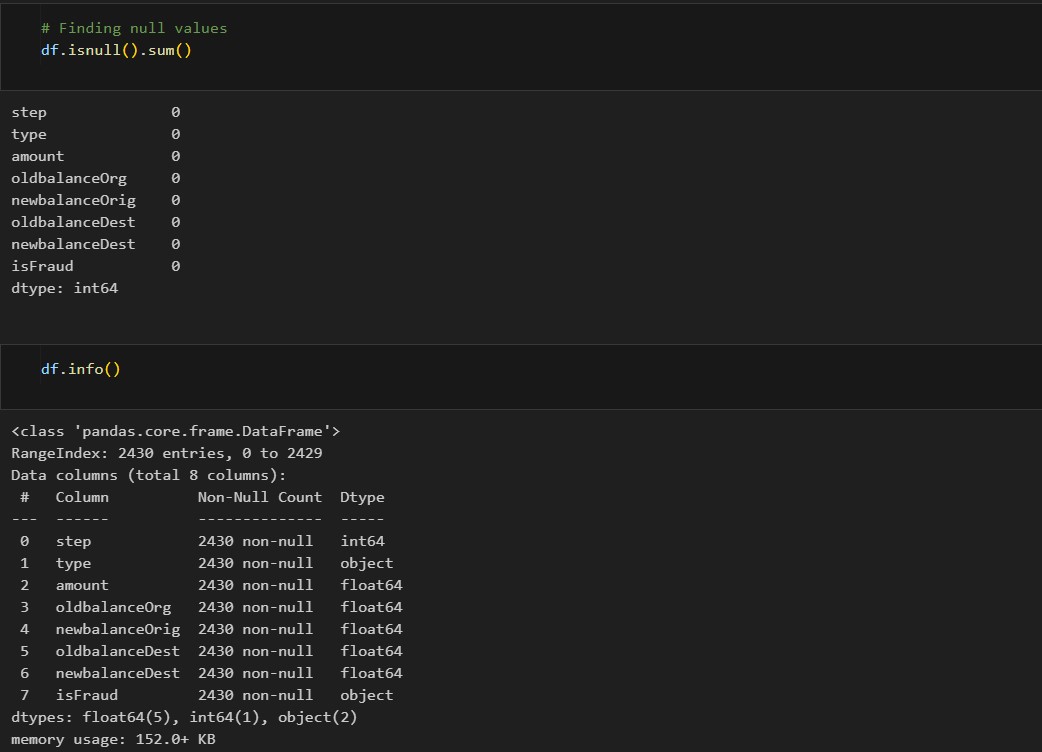


### 2.3 DATA PREPROCESSING



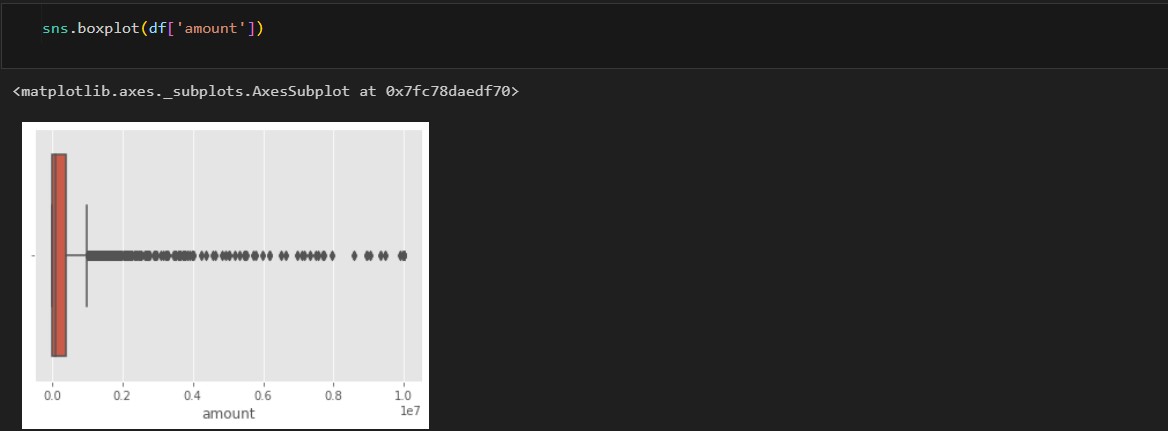
**FIGURE**

**Activity1: Checking null values**



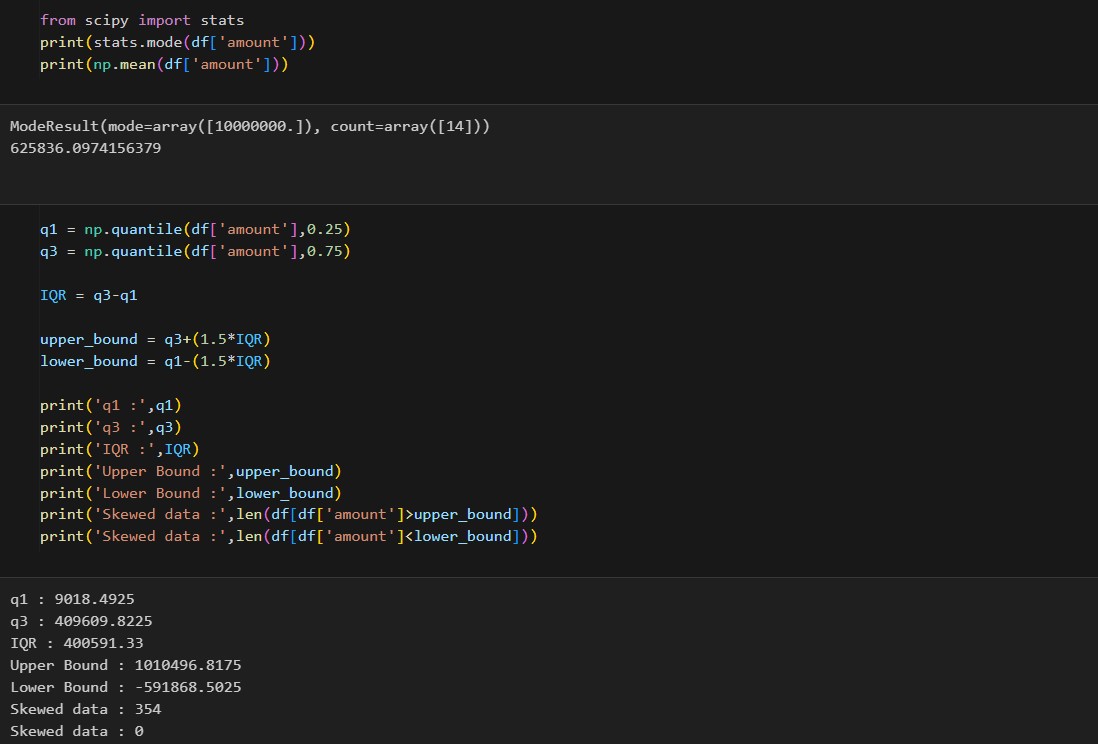
### FIGURE

#### Activity 2: Handling Outliers

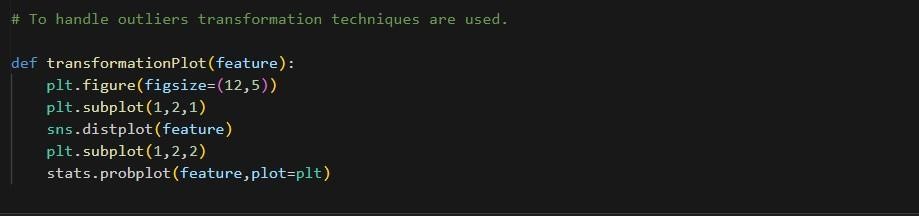


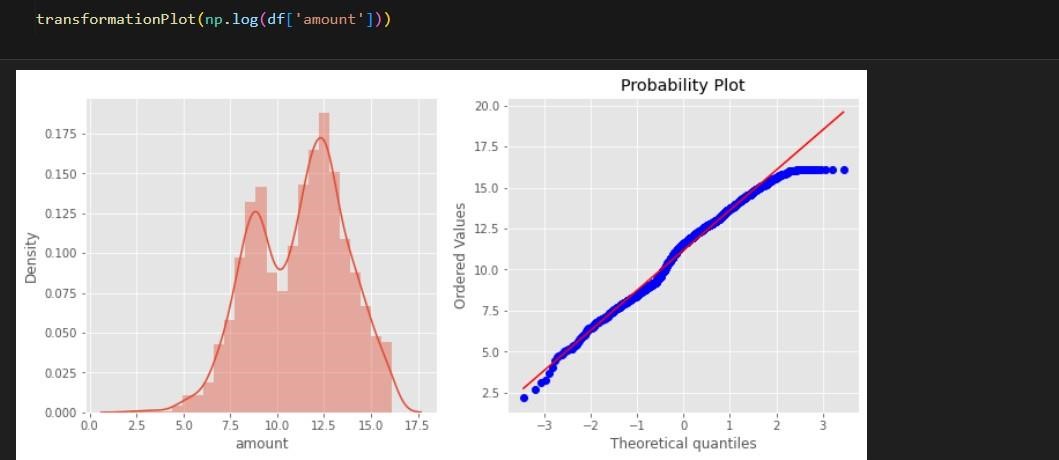
FIGURE

**Remove the outliers:**



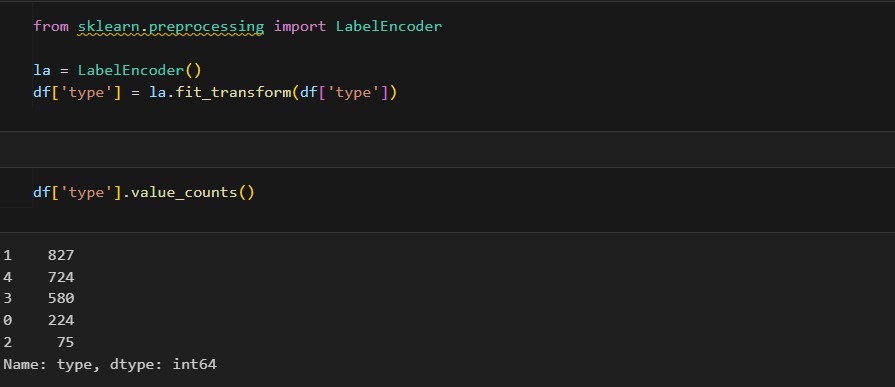
#### Figure





Fig

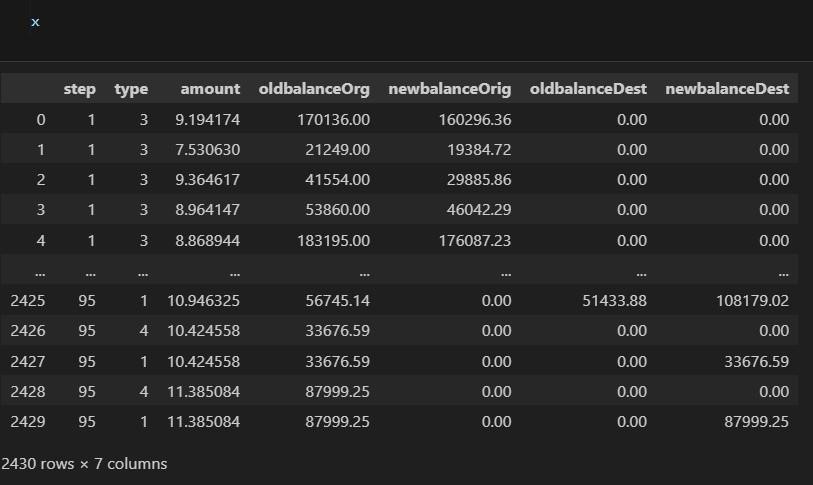
Activity 3:OBJECT DATA ENCODING



FIGURE

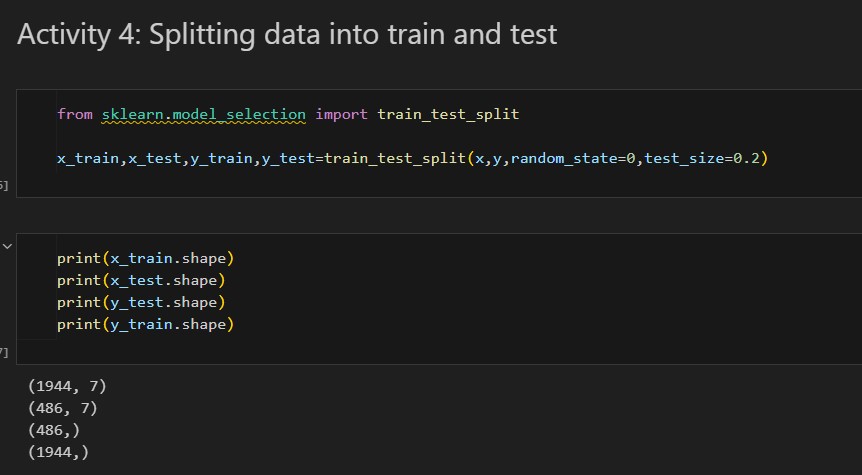
#### Divding data into dependent and independent





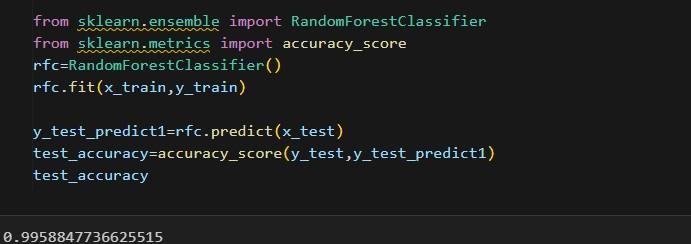
Figure

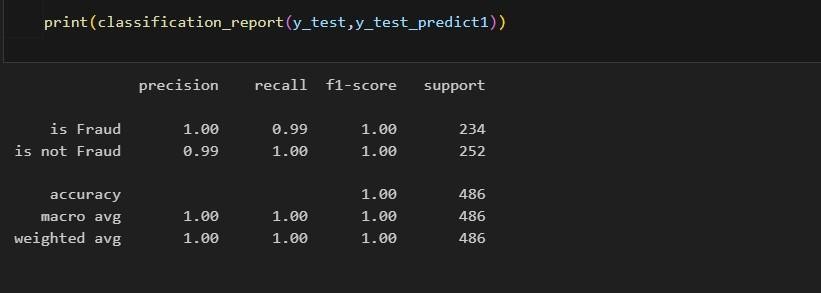




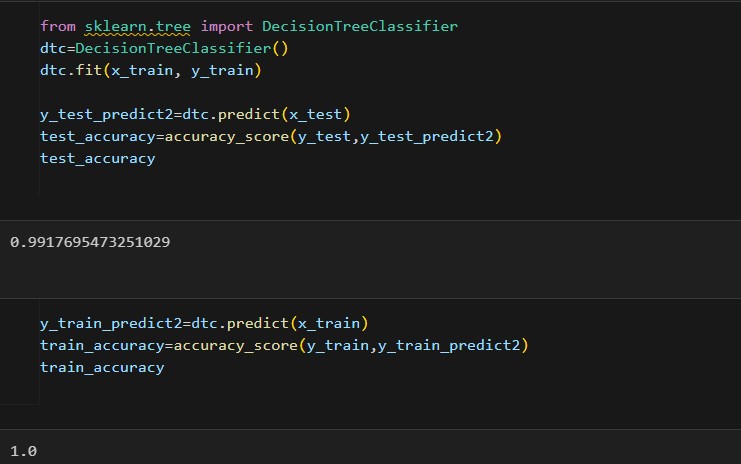
### MODEL BULDING

**Activity 1:Random forest classifier1**

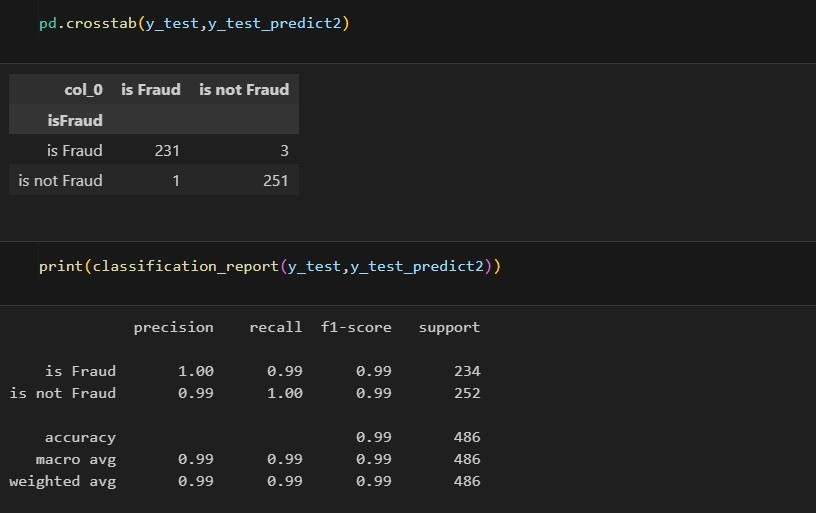




## Activity 2: Decision tree classifier

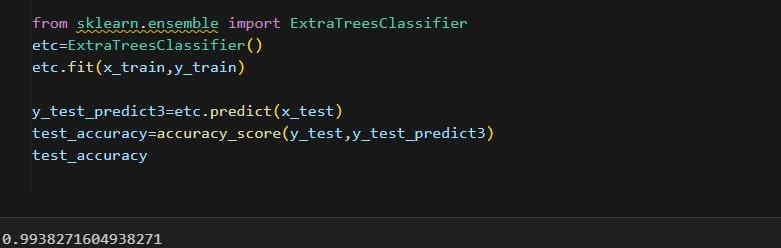


figure



figure

Activity 3



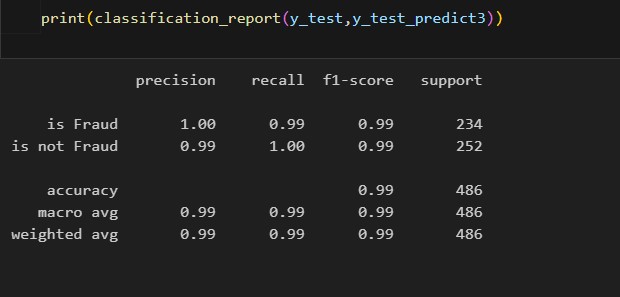
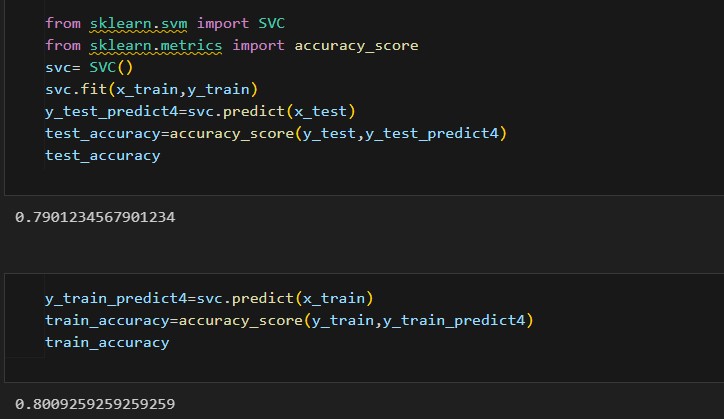
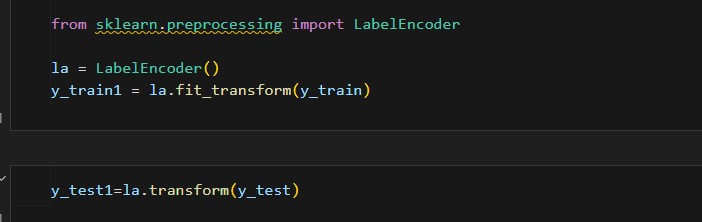
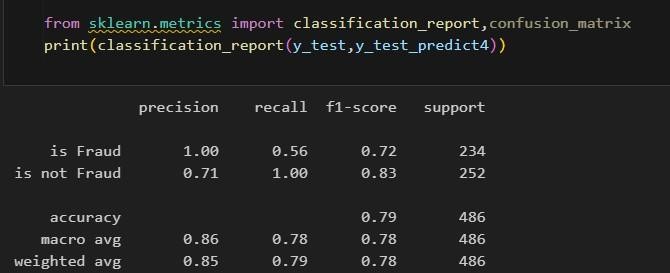
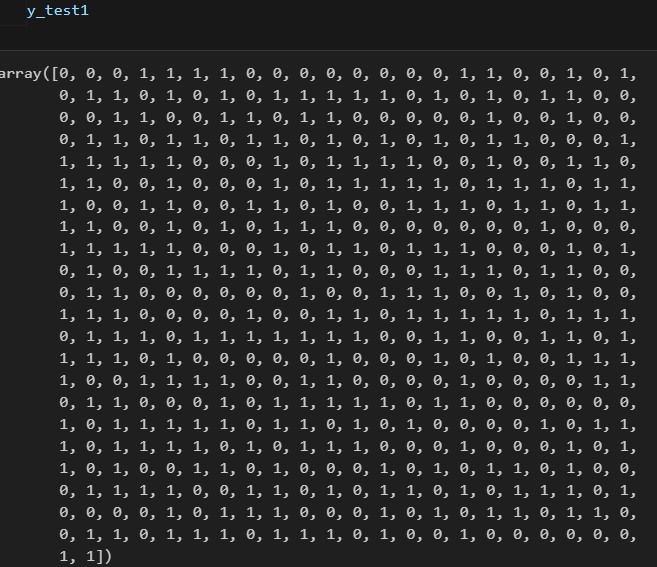


Figure Activity 4:

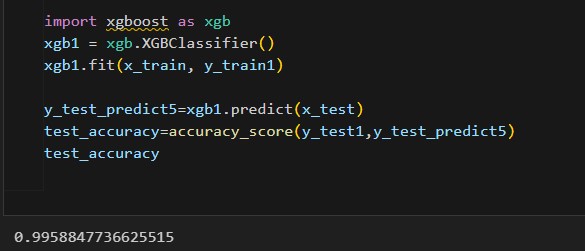
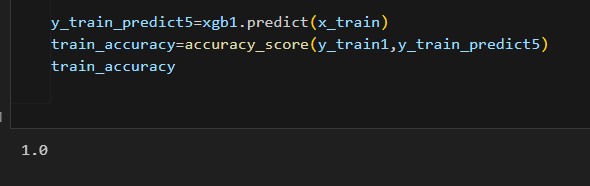


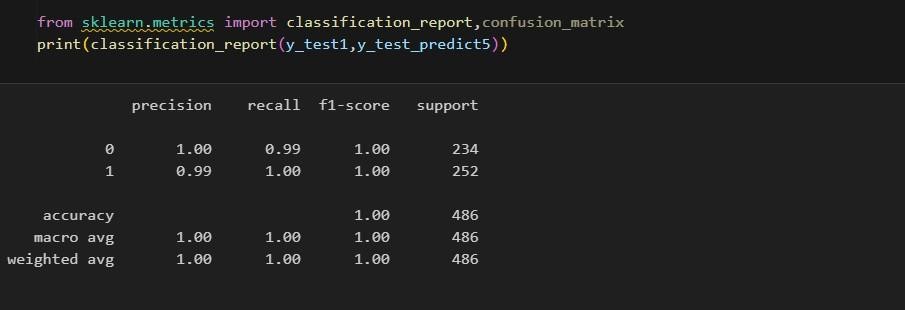




### Figure figu

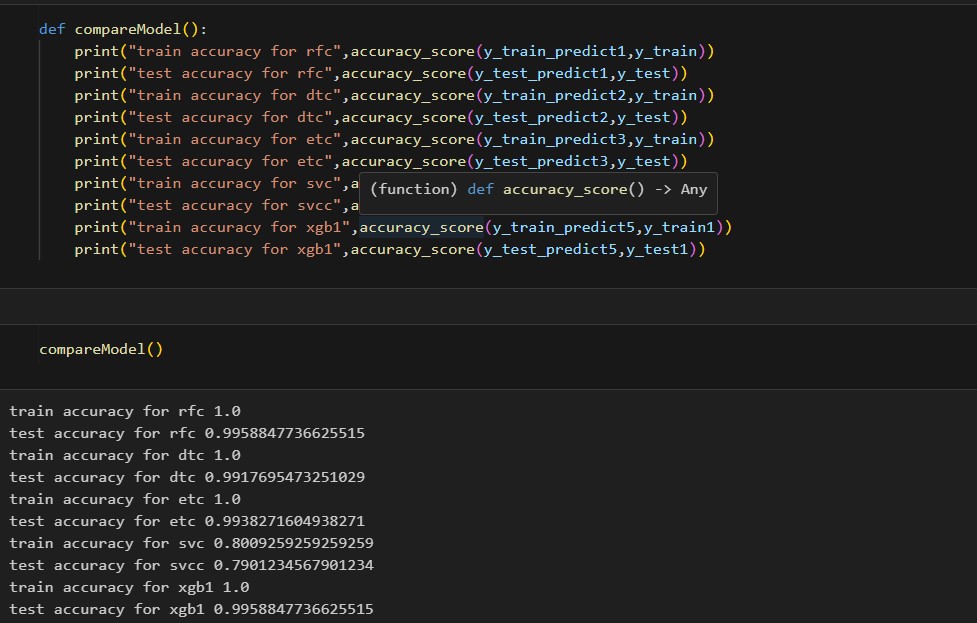
**Activity 5:**



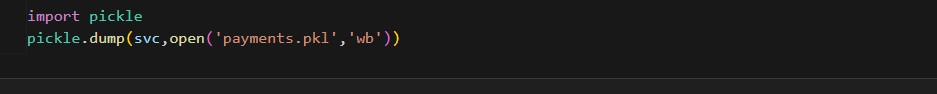


**Comparing The Models:**

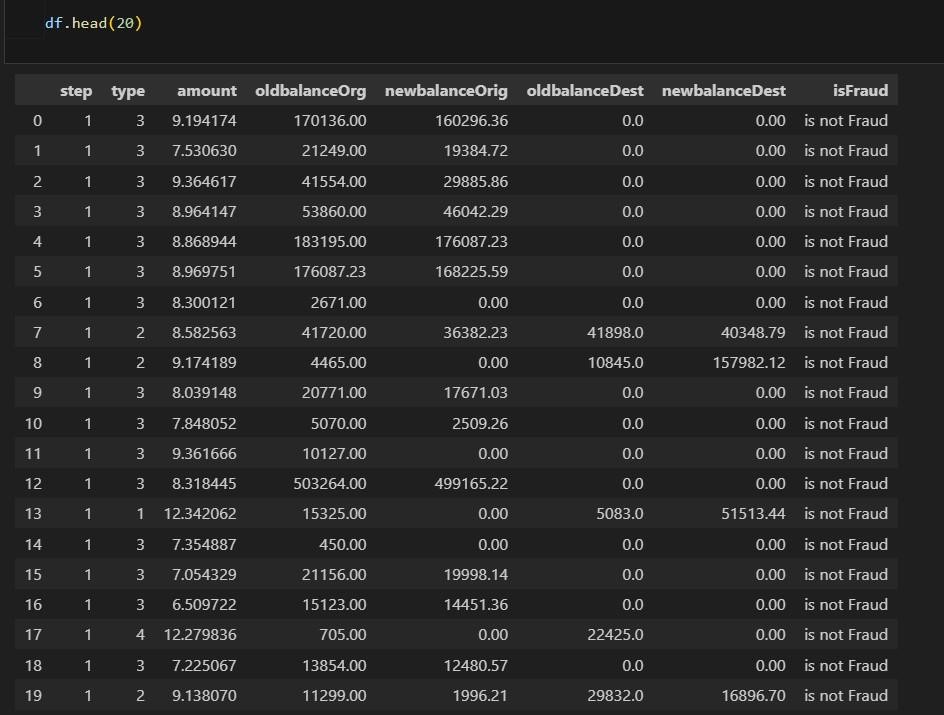
### figure

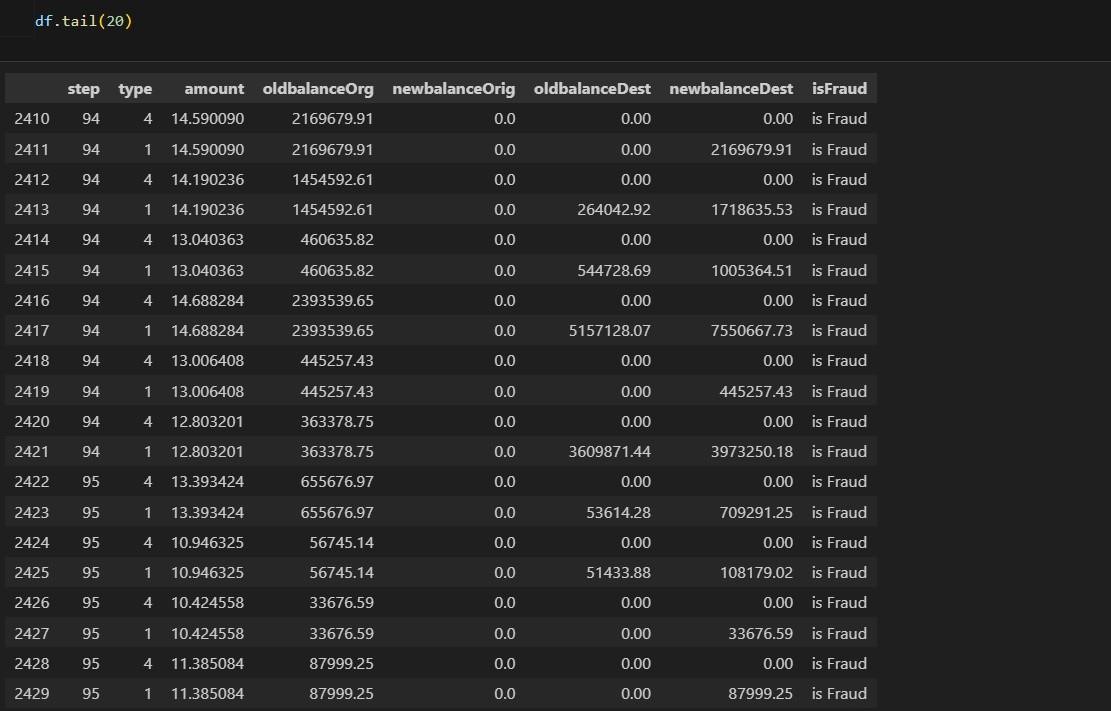


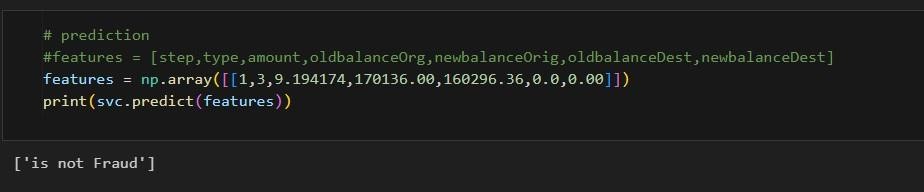
figure

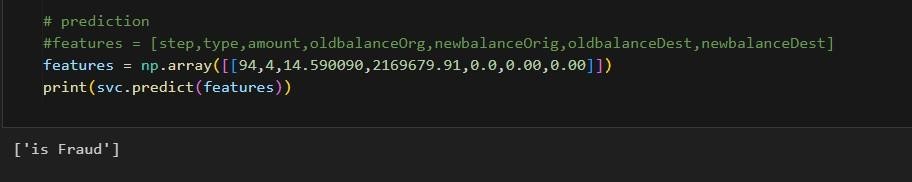


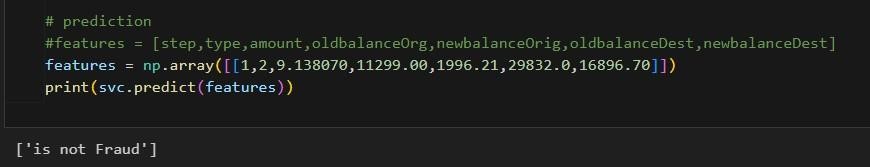
figure

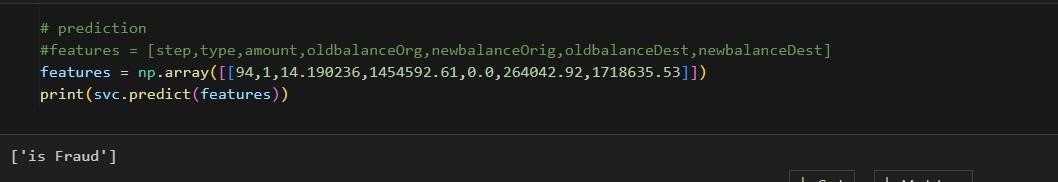






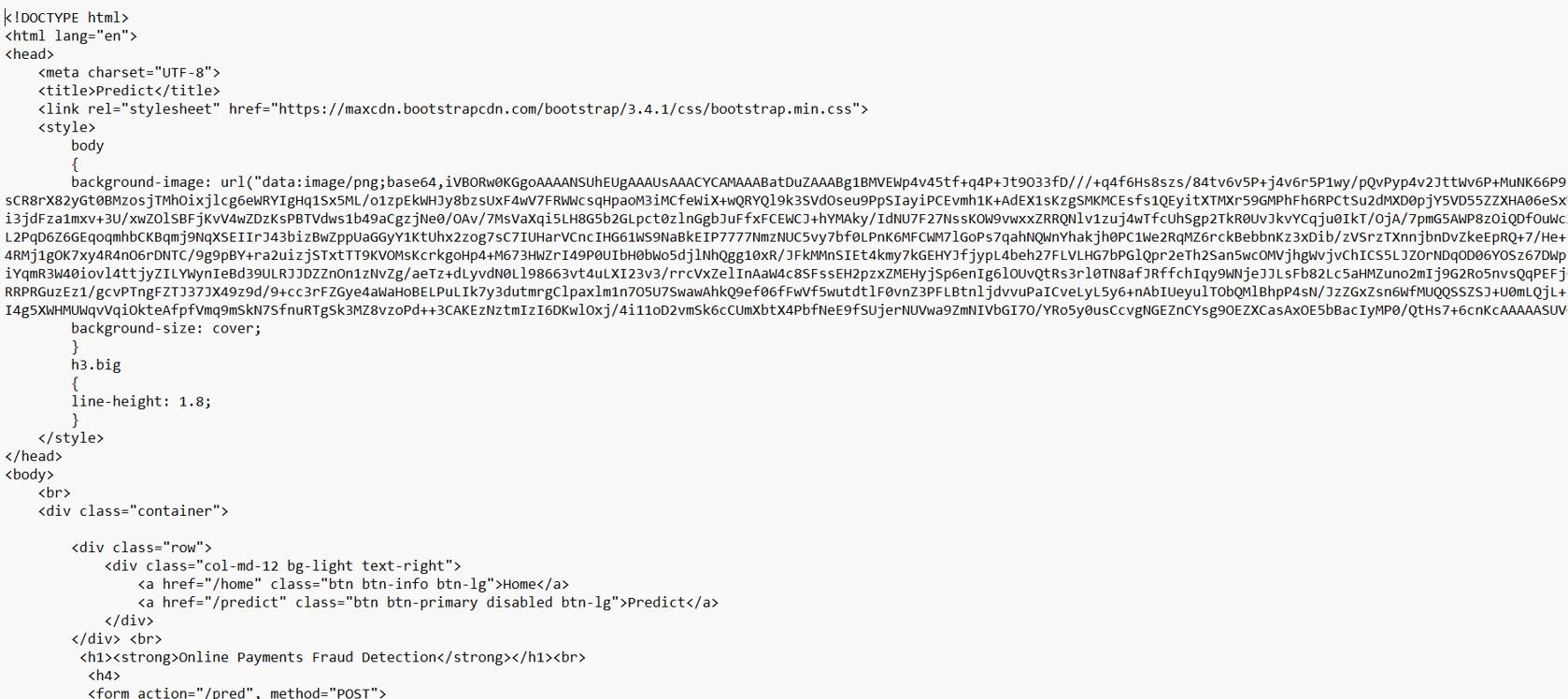


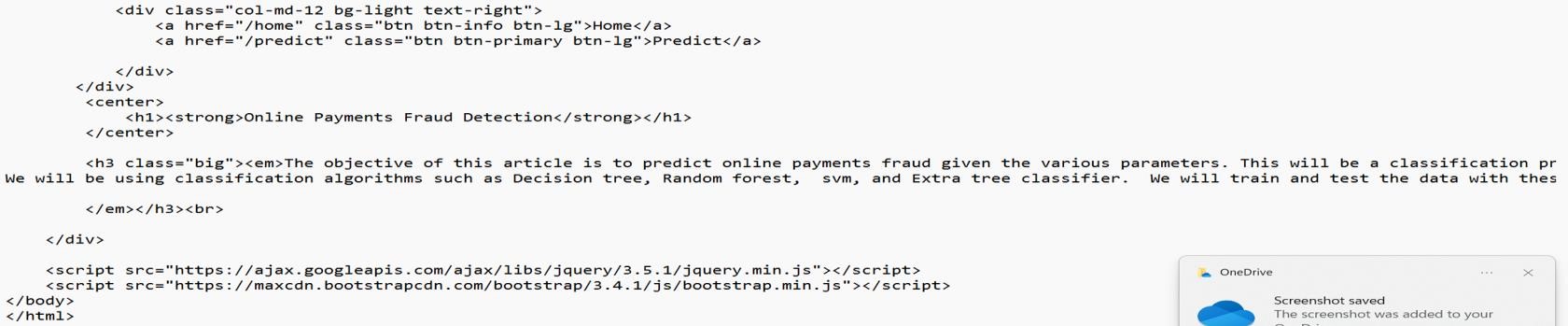
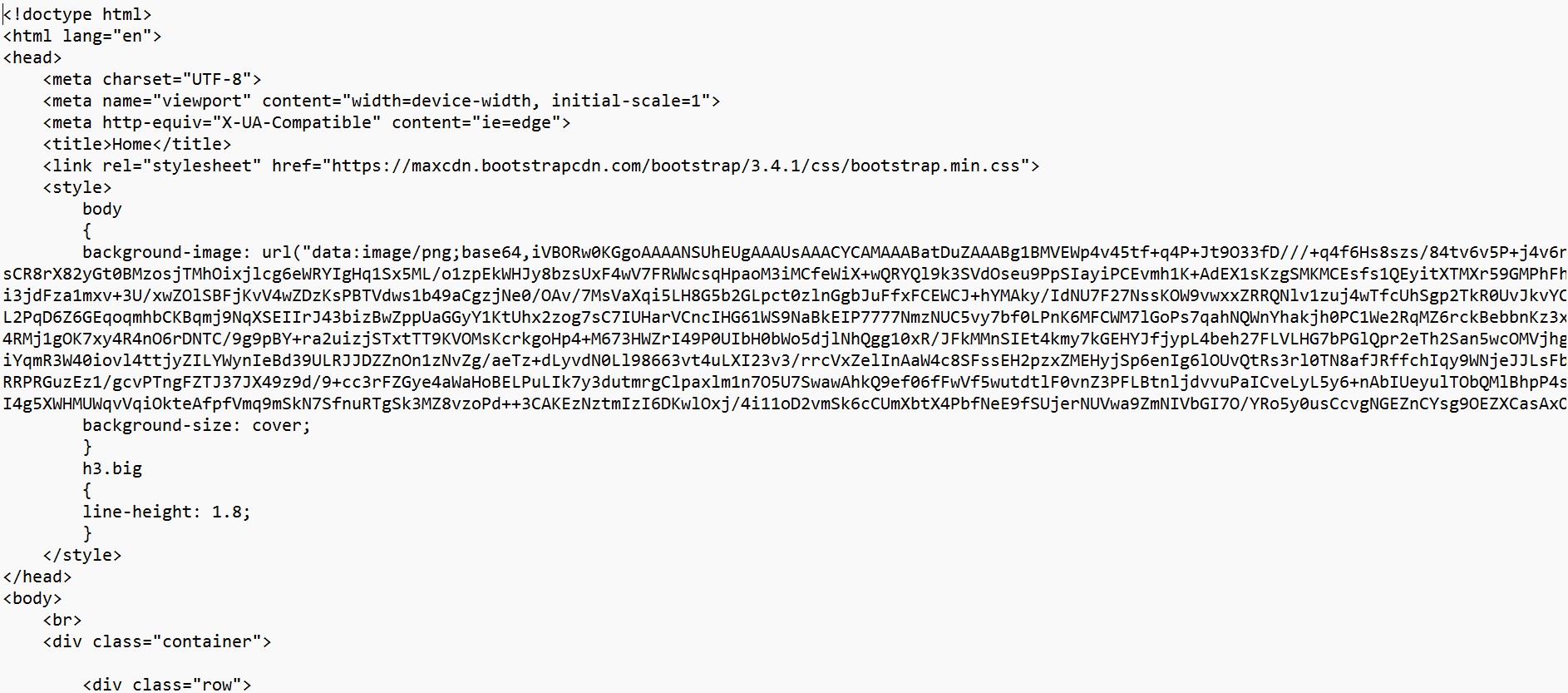




**2.1. APPLICATION BUILDING:**

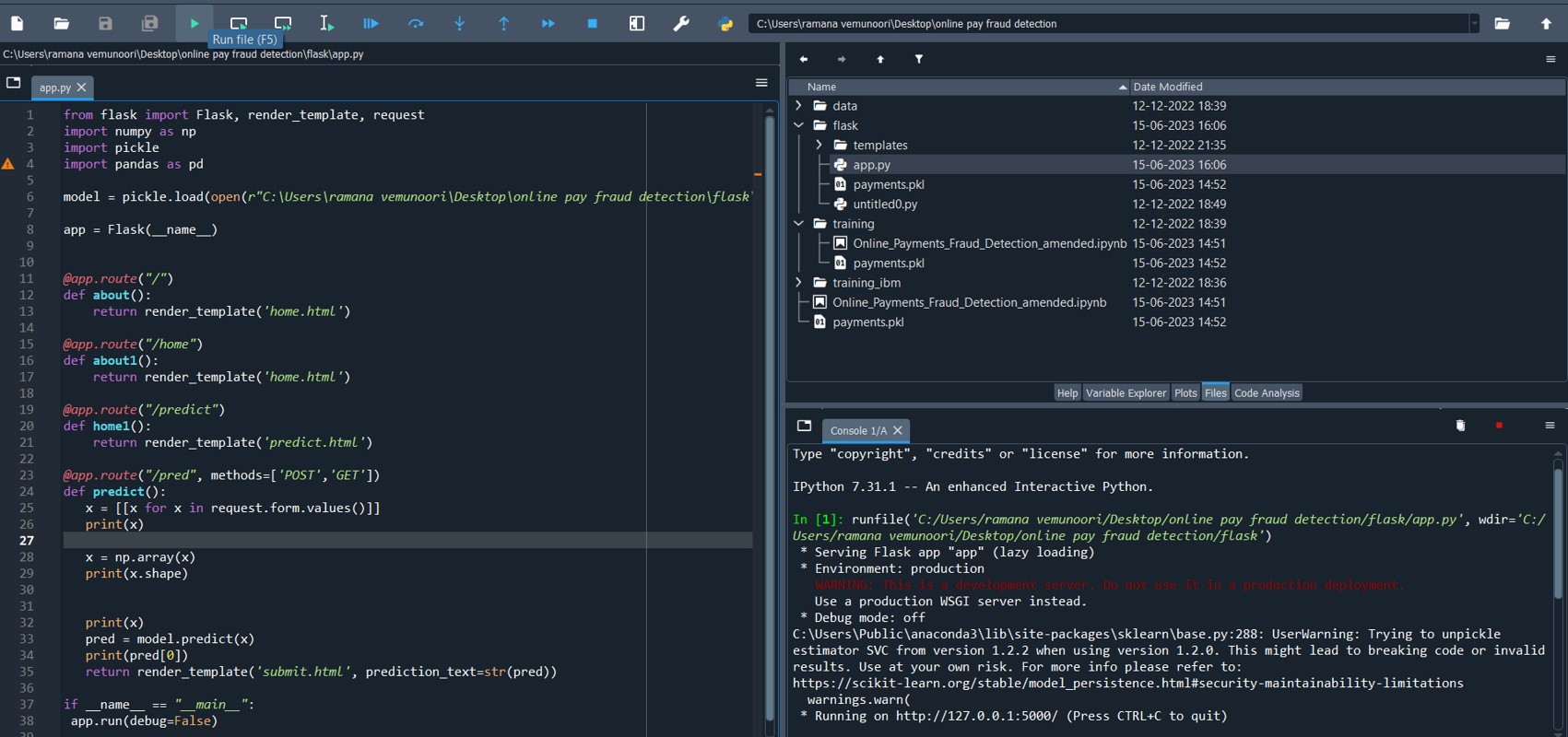
**HTML Code:**





#### Figure : html code



**Python code:**

**Web UI Page:**

**Figure**

### CONCLUSION

#### Inputs given to web page

 Figu



figure

# 3. APPLICATIONS

**The areas where this solution can be applied:**

o Can be applied in each and every individual’s Daily Life. o Bank transfers o Digital wallets like google pay o QR codes/UPI o BNPL

# 4. ADVANTAGES

**Some advantages of online payments:**

* Speed of transactions
* Convenience
* Reaching global audience
* Availability of more distribution channels
* Better customer experience
* Easy management
* Recurring payment capabilities
* Low transaction costs
* Quick and easy setup

# 5.DISADVANTAGES

**Some disadvantages of online payments:**

* Technical problems
* Password threats
* Cost of fraud
* Security Concerns ➢ False identity
* Loss of smart card
* Limitations on amount and time
* Service fees and other additional costs

# 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 Fraud Transaction Detection System is basically an extension of the existing system .Using This system, the algorithms will be built to through 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.

# 5. BIBILOGRAPHY

1. K.Chaudhary, J.Yadav, “A review of fraud: A comparative study.”decis. Support syst, vol 50, no3, pp.602-613,2011

1. 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 58, 2017.

1. Customer Transaction Fraud Detection Using Xgboost Model -by Yixuan Zhang, Ziyi Wang, Jialiang Tong, Fengqiang Gao June, 2020

1. Wang, M., Yu, J., & Ji, Z. (2018). Credit Fraud Risk Detection Based on XGBoost-LR Hybrid Model

1. 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

# HELP FILE

**PROJECT EXECUTION:**

**STEP-1:** Go to **Start,** search and launch **ANACONDA NAVIGATOR.**

**STEP-2:** After launching of **ANACONDA NAVIGATOR,** launch **JUPYTER NOTEBOOK.**

**STEP-3:** Open **“Major project code” 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, index.html, chance.html, nochance.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 all 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 ―**Go to Predict”.** Directly it will be navigated to index page.

**STEP-16:**A index page will be displayed where the user needs to give the inputs and then click on ―**Predict”**. Output will be generated whether a person is having liver disease or not.