



Department of Information Technology

NBA Accredited

A.P. Shah Institute of Technology

— G.B.Road,Kasarvadavli, Thane(W), Mumbai-400615

UNIVERSITY OF MUMBAI

Academic Year 2021-2022

A Project Report on

Credit Card fraud Detection

Submitted in partial fulfilment of the degree of
Bachelor of Engineering(Sem-8)

in

INFORMATION TECHNOLOGY

By

Sakshi Ravindra Shinde(19104055)

Divya Sanjiv Shepal (19104053)

Shubhangi Babaji Lanke (19104064)

Under the Guidance of

Name of Guide : Prof. Manjusha Kashilkar

1. Project Conception and Initiation

1.1 Objectives

- To Identify Fraudulent Credit card transaction
- To detect 100% of the Fraudulent transactions while minimizing the incorrect fraud classifications.
- To estimate, perceive or avoid objectionable behaviour , which consist of fraud , intrusion and defaulting .

1.2 Literature Review

- Fabiana Fournier, Ivo carriea, Inna skarbovsky, The Uncertain Case of Credit Card Fraud Detection, The 9th ACM International Conference On Distributed Event Based Systems(DEBS15) 2015
- A Comparative Analysis of Various Credit Card Fraud Detection Techniques:
- Iwasokun GB, Omomule TG, Akinyede RO. Encryption and tokenization-based system for credit card information security. Int J Cyber Sec Digital Forensics. 2018;7(3):283–93.

1.3 Problem Definition

- Unauthorized transactions and unwanted usage of an account by someone other than the owner of the account.
- Massive stream of payment requests were generated on daily basis.

1.4 Scope

- Can be applied in banking organizations.
- Can be used to decrease the fraud attempts.
- Can be used to enhance the process and guarantee the security of the account.

1.5 Technology stack

- VS Code
lightweight that provides the freedom to customize.

2. Project Design

2.1 Proposed System

- To help us classify fraudulent and legitimate credit card transaction by supervised learning Algorithm such as Random forest.
- To help us to get awareness about the fraudulent and without loss of any financially.

2.2 Design(Flow Of Modules)

- Importing Dataset
- Conversion of data into frames format
- Data Sampling
- Data Separation For Training and Testing
- Assigning train datasets to the models
- Applying algorithm and Creating the model
- Making Predictions for test dataset
- Calculating Accuracy

3. Implementation

File Edit View Navigate Code Refactor Run Tools VCS Window Help credit_card_fraud.py - credit_card_fraud.py

credit_card_fraud.py

Project

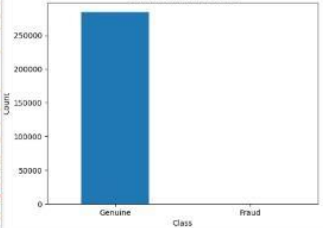
- credit_card_fraud.py
 - credit_card_fraud.py
 - metrics(actuals, predictions)
 - plot_confusion_matrix(cm, classes, ...)
 - External Libraries
 - Scratches and Consoles

credit_card_fraud.py

```
1 #!/usr/bin/env python
2 # coding: utf-8
3
4 ### Credit Card Fraud Detection Project
5
6 # In[17]:
7
8
9 # Import the necessary modules
10
11 import ...
12
13
14 # In[2]:
```

SciView: Data Plots

Visualization of Labels



| Class | Count |
|---------|----------|
| Genuine | ~240,000 |
| Fraud | ~400 |

Run: credit_card_fraud

```
29 Amount 284807 non-null float64
30 Class 284807 non-null int64
dtypes: float64(30), int64(1)
memory usage: 67.4 MB
Number of Genuine transactions: 284315
Number of Fraud transactions: 492
Percentage of Fraud transactions: 0.1727
Shape of train_X: (199364, 29)
Shape of test_X: (85443, 29)
```

Run TODO Problems Terminal Python Packages Python Console 1 Event Log

Python 3.8 has been configured as a project interpreter // Configure a Python interpreter... (a minute ago) 1:1 LF UTF-8 4 spaces Python 3.8

File Edit View Navigate Code Refactor Run Tools VCS Window Help credit_card_fraud.py - credit_card_fraud.py

myplot.png credit_card_fraud

Project credit_card_fraud.py

SciView: Data Plots

Run: credit_card_fraud

```
C:\Users\HP\AppData\Local\Programs\Python\Python38-32\python.exe C:/Users/HP/Downloads/credit-card-fraud-python-ml-code/credit_card_fraud.py
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 284807 entries, 0 to 284806
Data columns (total 31 columns):
#   Column   Non-Null Count  Dtype
---  ---
0    Time    284807 non-null  float64
1    V1       284807 non-null  float64
2    V2       284807 non-null  float64
3    V3       284807 non-null  float64
4    V4       284807 non-null  float64
5    V5       284807 non-null  float64
6    V6       284807 non-null  float64
7    V7       284807 non-null  float64
8    V8       284807 non-null  float64
9    V9       284807 non-null  float64
10   V10      284807 non-null  float64
11   V11      284807 non-null  float64
12   V12      284807 non-null  float64
13   V13      284807 non-null  float64
14   V14      284807 non-null  float64
15   V15      284807 non-null  float64
16   V16      284807 non-null  float64
```

Run TODO Problems Terminal Python Packages Python Console Event Log

Python 3.8 has been configured as a project interpreter // Configure a Python interpreter... (15 minutes ago) 7:38 LF UTF-8 4 spaces Python 3.8

File Edit View Navigate Code Refactor Run Tools VCS Window Help

credit_card_fraud.py - credit_card_fraud.py

myplot.png

Project

credit_card_fraud.py

280 # In[36]:

281

SciView: Data Plots

1 1:1

oit color) 15.91 kB

Database

SciView

Run: credit_card_fraud

14 V14 284807 non-null float64

15 V15 284807 non-null float64

16 V16 284807 non-null float64

17 V17 284807 non-null float64

18 V18 284807 non-null float64

19 V19 284807 non-null float64

20 V20 284807 non-null float64

21 V21 284807 non-null float64

22 V22 284807 non-null float64

23 V23 284807 non-null float64

24 V24 284807 non-null float64

25 V25 284807 non-null float64

26 V26 284807 non-null float64

27 V27 284807 non-null float64

28 V28 284807 non-null float64

29 Amount 284807 non-null float64

30 Class 284807 non-null int64

dtypes: float64(30), int64(1)

memory usage: 67.4 MB

Number of Genuine transactions: 284315

Number of Fraud transactions: 492

Percentage of Fraud transactions: 0.1727

Shape of train_X: (199364, 29)

Run

TODO

Problems

Terminal

Python Packages

Python Console

Event Log

Python 3.8 has been configured as a project interpreter // Configure a Python interpreter... (15 minutes ago)

7:38 LF UTF-8 4 spaces Python 3.8

File Edit View Navigate Code Refactor Run Tools VCS Window Help

credit_card_fraud.py - credit_card_fraud.py

myplot.png

credit_card_fraud

Project

Project

credit_card_fraud.py

credit_card_fraud.py

SciView: Data Plots

bit color) 15.91 kB

Run: credit_card_fraud

Number of Fraud transactions: 492
Percentage of Fraud transactions: 0.1727
Shape of train_X: (199364, 29)
Shape of test_X: (85443, 29)
Random Forest Score: 99.9637185023934
Decision Tree Score: 99.9204147794436
Confusion Matrix - Decision Tree
[[85265 42]
 [26 110]]
Confusion matrix, without normalization
Confusion Matrix - Random Forest
[[85301 6]
 [25 111]]
Confusion matrix, without normalization
Evaluation of Decision Tree Model

Accuracy: 0.99920
Precision: 0.72368
Recall: 0.80882
F1-score: 0.76389
Evaluation of Random Forest Model

Accuracy: 0.99964

Run

TODO

Problems

Terminal

Python Packages

Python Console

Event Log

Python 3.8 has been configured as a project interpreter // Configure a Python interpreter... (16 minutes ago)

7:38 LF UTF-8 4 spaces Python 3.8

File Edit View Navigate Code Refactor Run Tools VCS Window Help credit_card_fraud.py - credit_card_fraud.py

myplot.png

Project

credit_card_fraud.py

280 # In[36]:
281

SciView: Data Plots

Database

Run: credit_card_fraud

Confusion matrix, without normalization
Evaluation of Decision Tree Model

Accuracy: 0.99920
Precision: 0.72368
Recall: 0.80882
F1-score: 0.76389

Evaluation of Random Forest Model

Accuracy: 0.99964
Precision: 0.94872
Recall: 0.81618
F1-score: 0.87747

Resampled shape of X: (568630, 29)
Resampled shape of Y: (568630,)
Counter({0: 284315, 1: 284315})

Structure

5. Result



credit_card_fraud



SciView:

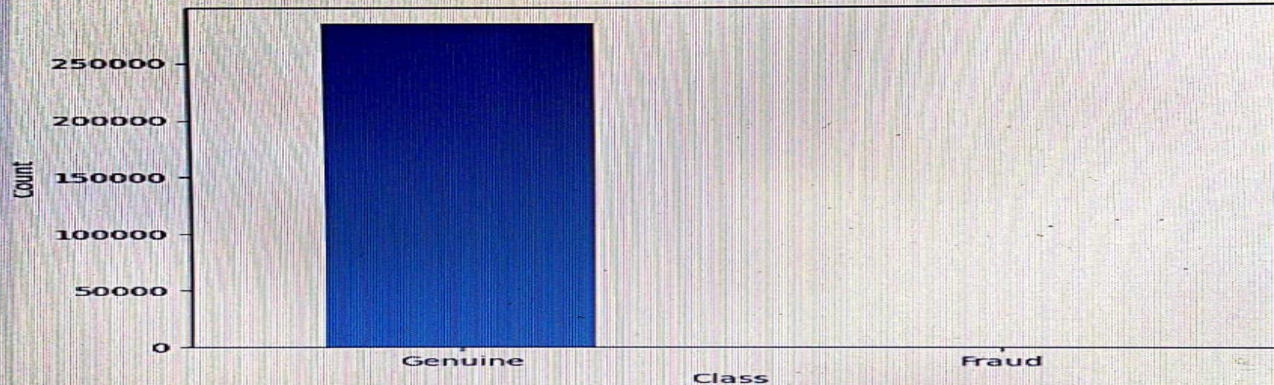
Data

Plots



bit color) 15.91 kB

Visualization of Labels



6. Conclusion and Future Scope

- A series of anti-fraud strategies can be adopted to prevent banks from great losses and reduce risks.
- This approach is ineffective in the long run as they are too broad to find ever more sophisticated forms of fraud.
- Due to the sensitive nature of the topic, often datasets are not available to effectively evaluate existing methods of fraud detection
- We cannot determine the names of fraud and genuine transactions for the given dataset using machine learning.

References

- <https://data-flair.training/blogs/credit-card-fraud-detection-python-machine-learning/amp/>
- <https://www.kaggle.com/code/fernolf/credit-card-fraud-detection/notebook>
- <http://sdiwc.net/digital-library/encryption-and-tokenizationbased-system-for-credit-card-information-security.html>

Thank You

