

Plot No. 98 Cidco Sector 3, Postbox No 15 Airoli Navi Mumbai 400708

MINI PROJECT REPORT

ON

Credit Card Fraud Detection

SUBMITTED TO

Prof. Swati Narane

by

Students of Third Year Information Technology

Report Submitted in Fulfilment of requirement of University of Mumbai

in

Data Science Using Python Lab

April 2022

TABLE OF CONTENTS

Acknowledgement 3

Abstract 4

Introduction 5

Problem Definition 6

Literature Survey 7

Design/Algorithm 8

Results and Discussion 9

Conclusion 10

**ACKNOWLEDGEMENT**

We have taken efforts in this report. However, it would not have been possible without the kind support and help of many organizations and individuals. We would like to extend our sincere thanks to all of them.

We are highly indebted to Prof.Swati Narwane for her guidance and constant supervision as well as for providing necessary information regarding the report and also for the support in completing it.

We would like to express our gratitude towards our parents for their kind co-operation which helped us in completing this report. Our thanks and appreciation also go to those people who have willingly helped us with their abilities.

Yours faithfully

TE\_B3\_44\_Shivani Shetty

TE\_B3\_53\_Tanushree Singh

TE\_B3\_55\_Jyoti Some

**ABSTRACT**

Fraud detection is a set of activities that are taken to prevent money or property from being obtained through false pretences. Fraud can be committed in different ways and in many industries. Credit card frauds are easy and friendly targets. E-commerce and many other online sites have increased the online payment modes, increasing the risk for online frauds. Increase in fraud rates, researchers started using different machine learning methods to detect and analyse frauds in online transactions. Credit card fraud generally happens when the card was stolen for any of the unauthorized purposes or even when the fraudster uses the credit card information for his use. Lots of money are lost due to credit card fraud every year. There is a lack of research studies on analysing real-world credit card data owing to confidentiality issues. In this paper, machine learning algorithms are used to detect credit card fraud. To evaluate the model efficacy, a publicly available credit card data set is used. The System prediction level & accuracy of fraud detection is not 100 percent accurate, So there is a chance of getting fraud. Then, a real-world credit card data set from a financial institution is analysed. In addition, noise is added to the data samples to further assess the robustness of the algorithms. The experimental results positively indicate that the majority voting method achieves good accuracy rates in detecting fraud cases in credit cards.

**INTRODUCTION**

A credit card is a thin handy plastic card that contains identification information such as a signature or picture, and authorizes the person named on it to charge purchases or services to his account - charges for which he will be billed periodically. Today, the information on the card is read by automated teller machines (ATMs), store readers, bank and is also used in online internet banking system. They have a unique card number which is of utmost importance. Its security relies on the physical security of the plastic card as well as the privacy of the credit card number. There is a rapid growth in the number of credit card transactions which has led to a substantial rise in fraudulent activities. Credit card fraud is a wide-ranging term for theft and fraud committed using a credit card as a fraudulent source of funds in a given transaction. Meta learning and pattern matching. The Genetic algorithms are evolutionary algorithms which aim to obtain the better solutions in eliminating the fraud. A high importance is given to develop efficient and secure electronic payment system to detect whether a transaction is fraudulent or not. In this report, we will focus on credit card fraud and its detection measures. A credit card fraud occurs when one individual uses other individuals’ card for their personal use without the knowledge of its owner. When such kind of cases takes place by fraudsters, it is used until its entire available limit is depleted.

**Problem Definition**

The Credit Card Fraud Detection Problem includes modeling past credit card transactions with the knowledge of the ones that turned out to be fraud. This model is then used to identify whether a new transaction is fraudulent or not. Our aim here is to detect 100% of the fraudulent transactions while minimizing the incorrect fraud classifications. What does credit card fraud detection include?

The Credit Card Fraud Detection Problem includes modeling past credit card transactions with the knowledge of the ones that turned out to be fraud. This model is then used to identify whether a new transaction is fraudulent or not.

**Literature Survey**

Fraud act as the unlawful or criminal deception intended to result in financial or personal benefit. It is a deliberate act that is against the law, rule or policy with an aim to attain unauthorized financial benefit.

The fraud detection is a complex task and there is no system that correctly predicts any transaction as fraudulent. The properties for a good fraud detection system are:

1. Should identify the frauds accurately.

2. Should detect the frauds quickly.

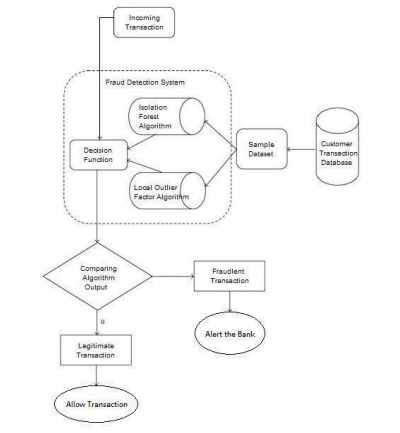
3. Should not classify a genuine transaction as fraud.

Outlier detection is a critical task as outliers indicate abnormal running conditions from which significant performance degradation may happen. Techniques used in fraud detection can be divided into two:

1) Supervised techniques where past known legitimate/fraud cases are used to build a model which will produce a suspicion score for the new transactions.

2) Unsupervised are those where there are no prior sets in which the state of the transactions are known to be fraud or legitimate

**Design/Algorithm**

****

Block Diagram (Credit Card Fraud Detection)

**Results**

In [1]:

**import** **sys**

**import** **numpy**

**import** **pandas**

**import** **matplotlib**

**import** **seaborn**

**import** **scipy**

**print**('Python: {}'.format(sys.version))

**print**('Numpy: {}'.format(numpy.\_\_version\_\_))

**print**('Pandas: {}'.format(pandas.\_\_version\_\_))

**print**('Matplotlib: {}'.format(matplotlib.\_\_version\_\_))

**print**('Seaborn: {}'.format(seaborn.\_\_version\_\_))

**print**('Scipy: {}'.format(scipy.\_\_version\_\_))

In [2]:

*# import the necessary packages*

**import** **numpy** **as** **np**

**import** **pandas** **as** **pd**

**import** **matplotlib.pyplot** **as** **plt**

**import** **seaborn** **as** **sns**

In [11]:

**from** **sklearn.metrics** **import** classification\_report, accuracy\_score

**from** **sklearn.ensemble** **import** IsolationForest

**from** **sklearn.neighbors** **import** LocalOutlierFactor

*# define random states*

state = 1

*# define outlier detection tools to be compared*

classifiers = {

"Isolation Forest": IsolationForest(max\_samples=len(X),

contamination=outlier\_fraction,

random\_state=state),

"Local Outlier Factor": LocalOutlierFactor(

n\_neighbors=20,

contamination=outlier\_fraction)}

In [15]:

*# Fit the model*

plt.figure(figsize=(9, 7))

n\_outliers = len(Fraud)

**for** i, (clf\_name, clf) **in** enumerate(classifiers.items()):

*# fit the data and tag outliers*

**if** clf\_name == "Local Outlier Factor":

y\_pred = clf.fit\_predict(X)

scores\_pred = clf.negative\_outlier\_factor\_

**else**:

clf.fit(X)

scores\_pred = clf.decision\_function(X)

y\_pred = clf.predict(X)

*# Reshape the prediction values to 0 for valid, 1 for fraud.*

y\_pred[y\_pred == 1] = 0

y\_pred[y\_pred == -1] = 1

n\_errors = (y\_pred != Y).sum()

*# Run classification metrics*

**print**('{}: {}'.format(clf\_name, n\_errors))

**print**(accuracy\_score(Y, y\_pred))

**print**(classification\_report(Y, y\_pred))

**Conclusion**

Credit card fraud is without a doubt an act of criminal dishonesty. This report has listed out the most common methods of fraud along with their detection methods and reviewed recent findings in this field. This report has also explained in detail, how machine learning can be applied to get better results in fraud detection along with the algorithm, pseudocode, explanation its implementation and experimentation results. If this project were to be used on a commercial scale. Being based on machine learning algorithms, the program will only increase its efficiency over time as more data is put into it.

# REFERENCES

* <https://www.python.org/>
* <https://pandas.pydata.org/>
* <https://numpy.org/>
* [ResearchGate](https://www.researchgate.net/publication/336800562_Credit_Card_Fraud_Detection_using_Machine_Learning_and_Data_Science/link/5db2f6dd92851c577ec2e973/download)