

Credit Card Fraud Detection using Machine Learning Algorithms

By

Sai Gogul Krishnan BK

Sudharson P

Jamal Mohideen S

INTRODUCTION

Credit card is a small thin plastic or fiber card that contains information about the person such as picture or signature and person named on it to charge purchases and service to his linked account charges for which will be debited regularly. Now a day's card information is read by ATM's, swiping machines, store readers, bank and online transaction. Each card as a unique card number which is very important, its security is mainly relies on physical security of the card and also privacy of the credit card number.

PROBLEM STATEMENT

Credit card frauds are increasing heavily because of fraud financial loss is increasing drastically. Every year due to fraud Billions of amounts lost. To analyze the fraud

there is lack of research. Many machine learning algorithms are implemented to detect real world credit card fraud. ANN and hybrid algorithms are applied.

OBJECTIVES

The objectives of the project is to implement machine learning algorithms to detect credit card fraud detection with respect to time and amount of transaction.

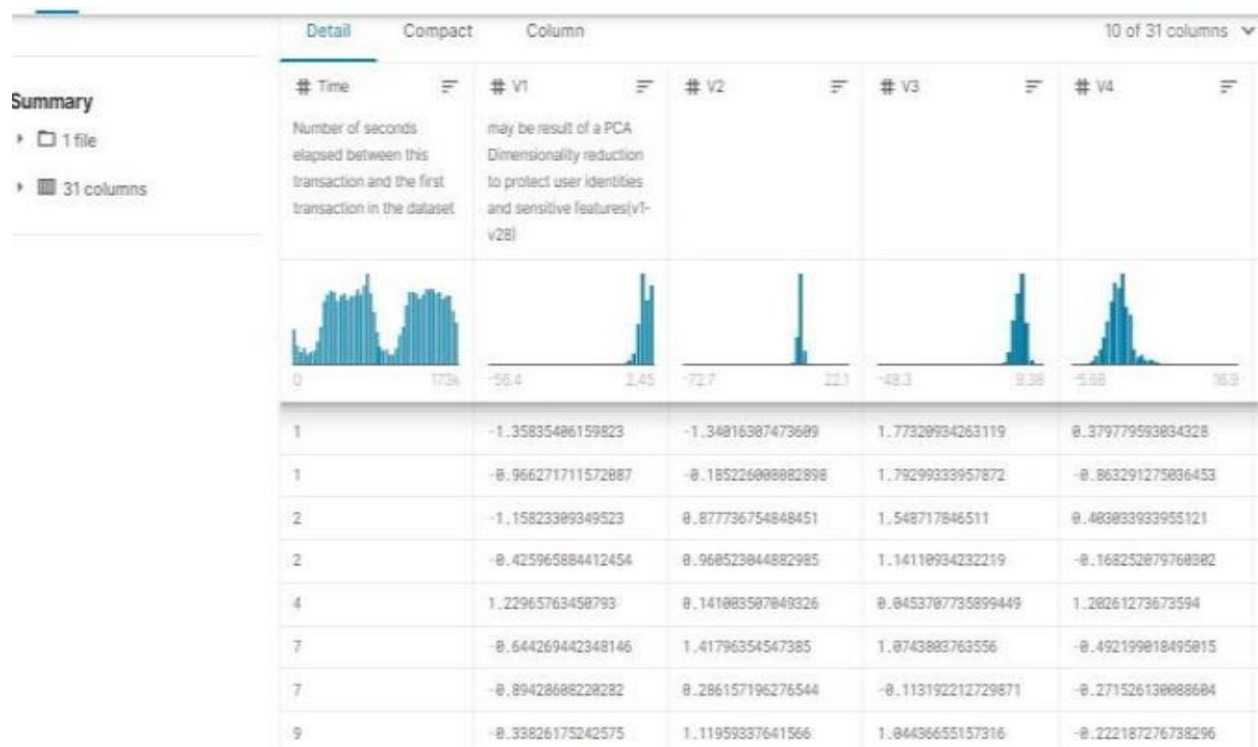
RELATED WORKS

In previous studies, many methods have been implemented to detect fraud using supervised, unsupervised algorithms and hybrid ones. Fraud types and patterns are evolving day by day. It is important to have clear understanding of technologies behind fraud detection. Here discuss machine learning models, algorithms and fraud detection models used in earlier studies. In data mining techniques are discussed and these methods take time dealing with huge data. Overlapping is another problem with credit card transaction data preparation. Imbalanced data distribution is overcome using sampling methods.

TRANSACTION DATABASE

Dataset contains the transaction from Europe card owners during September 2013. In this 492 out of

2,84,807 are fraud transactions. Data is not balanced because less amount of fraud cases as compared to huge transaction data. Dataset is converted PCA transformation and contains only numeric values. Due to privacy and confidentiality many background information is not provided only PCA transformed data is given. Only time and amount are not transformed to PCA all other given values v1, v2, v3 v28 are PCA transformed numeric values. Feature class contains 1 for fraud and 0 for normal transaction.

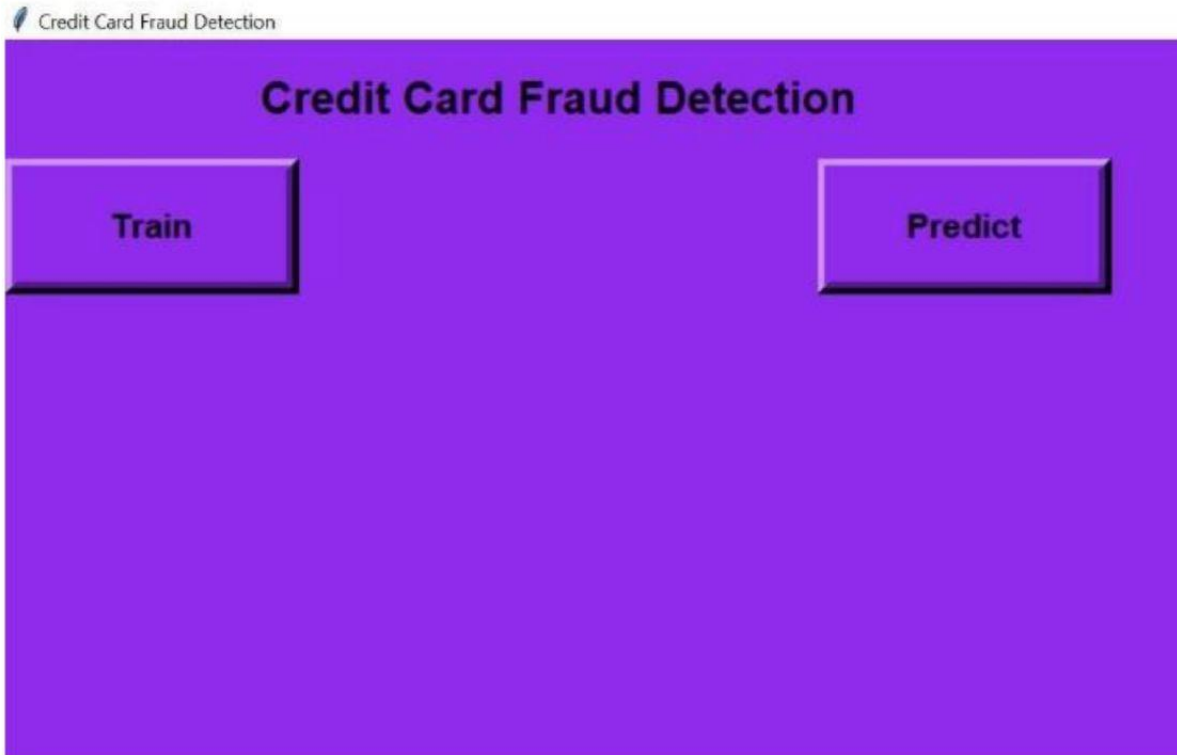


DESIGN AND IMPLEMENTATION OF ALGORITHMS

The procedure which we followed to predict the result are understanding problem statement and data by performing statistical analysis and visualization then checking whether the data is balance or not, In this data set the data is imbalanced, balanced by using oversampling, then scaling the data using standardization and normalization and testing data with different ML algorithms For any data science project some package are very important such as Numpy that is numeric python And pandas and for visualization of the data, matplotlib and seaborn is used which build on matplotlib with some extra features

RESULTS AND DISCUSSION

the user interface for test and train the data. Train and Test buttons are given to the user where using train the algorithms are trained and then o predict the fraud by clicking predict button it will take to another window where the input is given and output is seen as fraud or nonfraud.



CONCLUSION

Credit card fraud is most common problem resulting in loss of lot money for peoples and loss for some banks and credit card company. This project want to help the peoples from their wealth loss and also for the banked company and trying to develop the model which more efficiently separate the fraud and fraud less transaction by using the time and amount feature in data set given in the Kegel. first we build the model using some machine learning algorithms such as logistic regression, decision tree, support vector machine, this all are supervised machine learning algorithm in machine learning.