

Credit Card Fraud Detection

Course : IT-563 Data Mining
Instructor: PM Jat

Kaushal Patel (201501219)

Introduction

- The credit card has become the most popular mode of payment for both online as well as regular purchase.
- How does the credit card fraud takes place?
 - Phishing
 - Skimming
- How credit card companies detect fraud?
 - Credit card companies look for the patterns and search for the anomaly.
- Fraudsters are so experts that they generate new ways of committing fraudulent transactions each day which demands constant innovation for its detection techniques.

Related Research Work

- Classification Techniques
 - Naive Bayes approach
 - Decision tree
 - K-nearest neighbours
 - Support vector machine
- Data set
 - Financial transactions
 - 1,00,000 instances
 - Legitimate: fraud transactions
 - 97:3 ratio of outliers

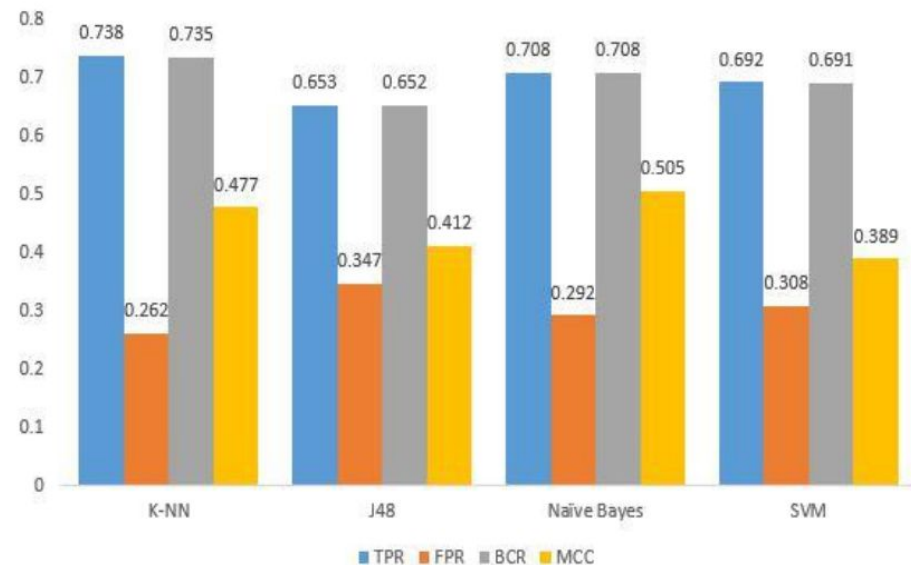
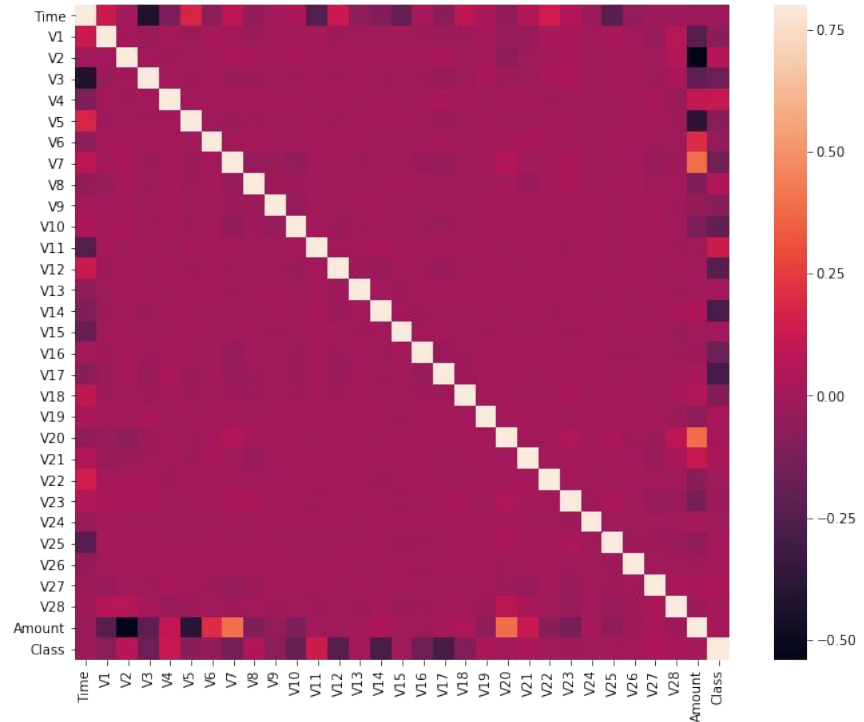
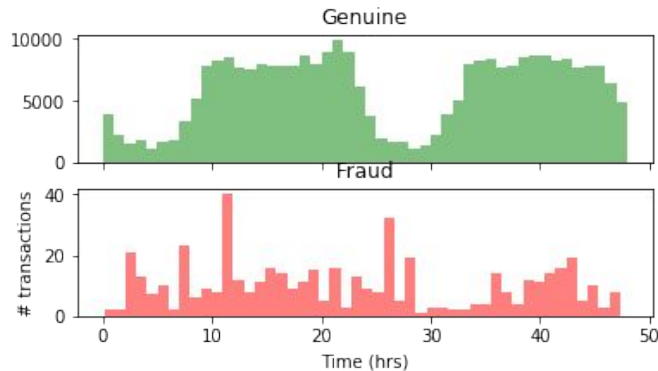


Figure 1. Performance of the four fraud detection models

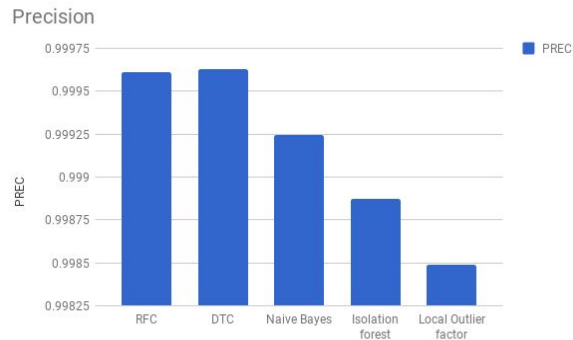
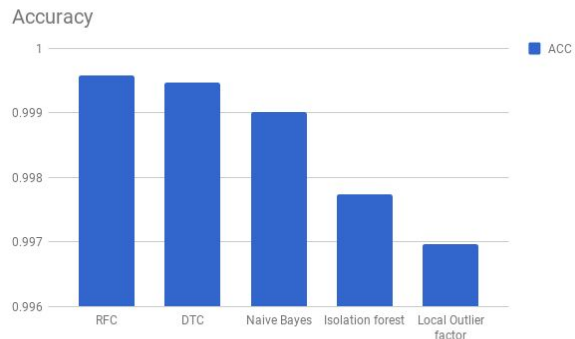
Data Set

- Transactions of european cardholders
- Total Transactions: 2,84,807
- Fraud Transactions: 492
- No of features :28



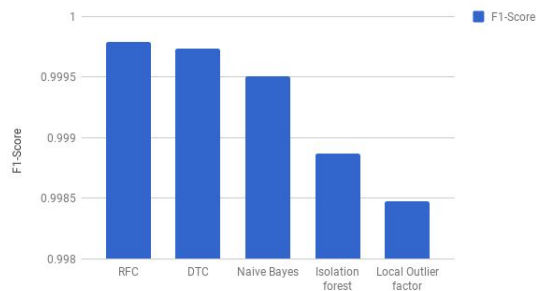
Classification Techniques

- There are five different classifier techniques used:
 - Random Forest Classifier
 - Decision Tree Classifier
 - Naive Bayes Classifier
 - Isolation forest
 - Local Outlier Factor
- Performance Metrics Used:
 - Accuracy
 - Precision
 - Recall
 - F1-score
 - False Positive Rate



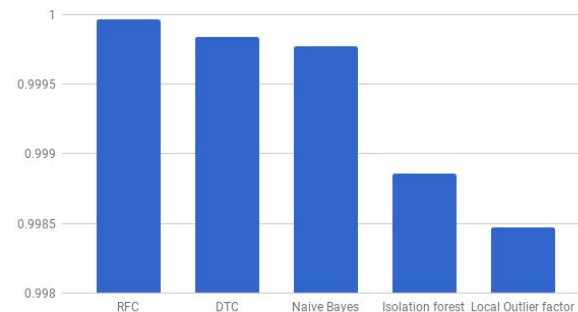
Experimental Results

F1-Score

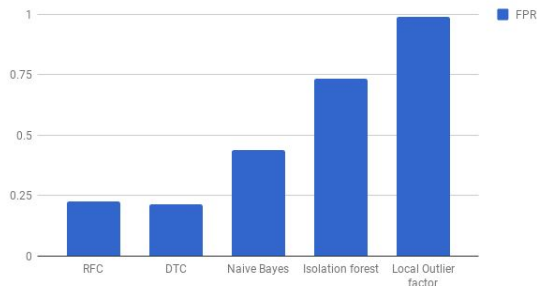


- False positive in this context is the fraud transactions which are detected legitimate.

Recall



False Positive Rate



Conclusions

- The results showed that there is no data mining technique that is universally better than others.
- Experimental Results depend on various factors:
 - Data instances
 - Outlier factor
 - No of features
 - Data preprocessing
- Performance improvement could be achieved through developing a fraud detection model using a combination of different data mining techniques (ensemble).

Take Away from this Project

- Various Python libraries :
 - Numpy
 - Pandas
 - Seaborn
 - Scipy
 - SKlearn
- Various fraud detection techniques and their usage.
- Various applications of Data mining to solve real life problems.

Any Questions ?

References :

[1] Data Mining Techniques for Credit Card Fraud Detection Empirical Study.pdf

[2] Credit card fraud and detection techniques: a review by Linda Delamaire (UK), Hussein Abdou (UK), John Pointon (UK)

[3] Survey on Credit Card Fraud Detection Techniques:

<http://www.ijecs.in/issue/v4i11/25%20ijecs.pdf>

[4] Credit Card Fraud Detection using Local Outlier Factor: by International Journal of Pure and Applied Mathematics

[5] <https://www.kaggle.com/samkirkiles/credit-card-fraud/data>

[6] <https://classeval.wordpress.com/introduction/basic-evaluation-measures/>



THANK YOU