Anomaly Detection in Payments

IE 406 : Machine Learning

Group no. 18

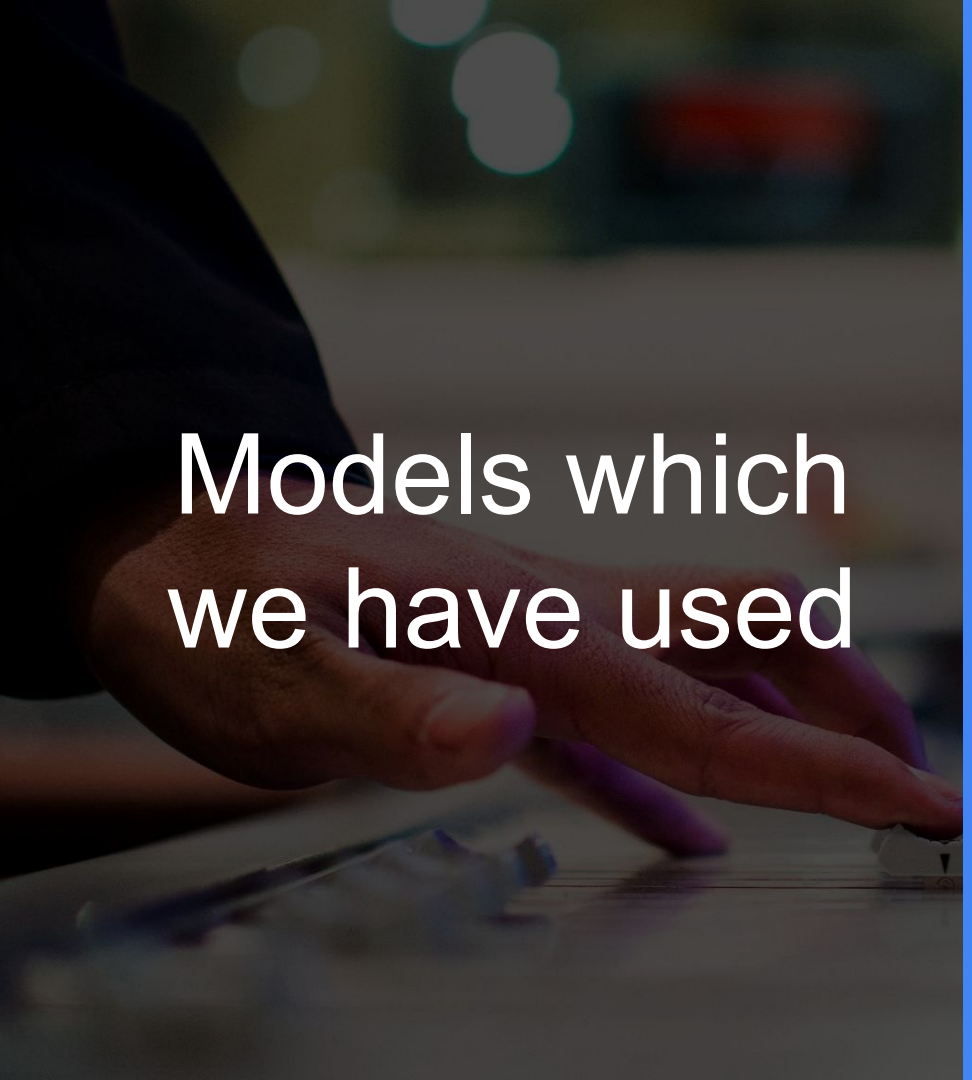
Assigned By

Prof. M.V. Joshi

Problem Statement

In training dataset, there are transactions made by credit cards in September 2013 by european cardholders. The goal is to separate fraudulent and normal transactions for which we have used two supervised and two unsupervised algorithms.

Github link : https://github.com/201701203/Anomaly-detection-in-Payments-using-ML



Models which we have used

1. Logistic Regression
2. SVM (kernel)
3. Isolation Forest
4. Local Outlier Factor (LOF)

Logistic Regression:-

Time :

Accuracy :

Support Vector Machine:-

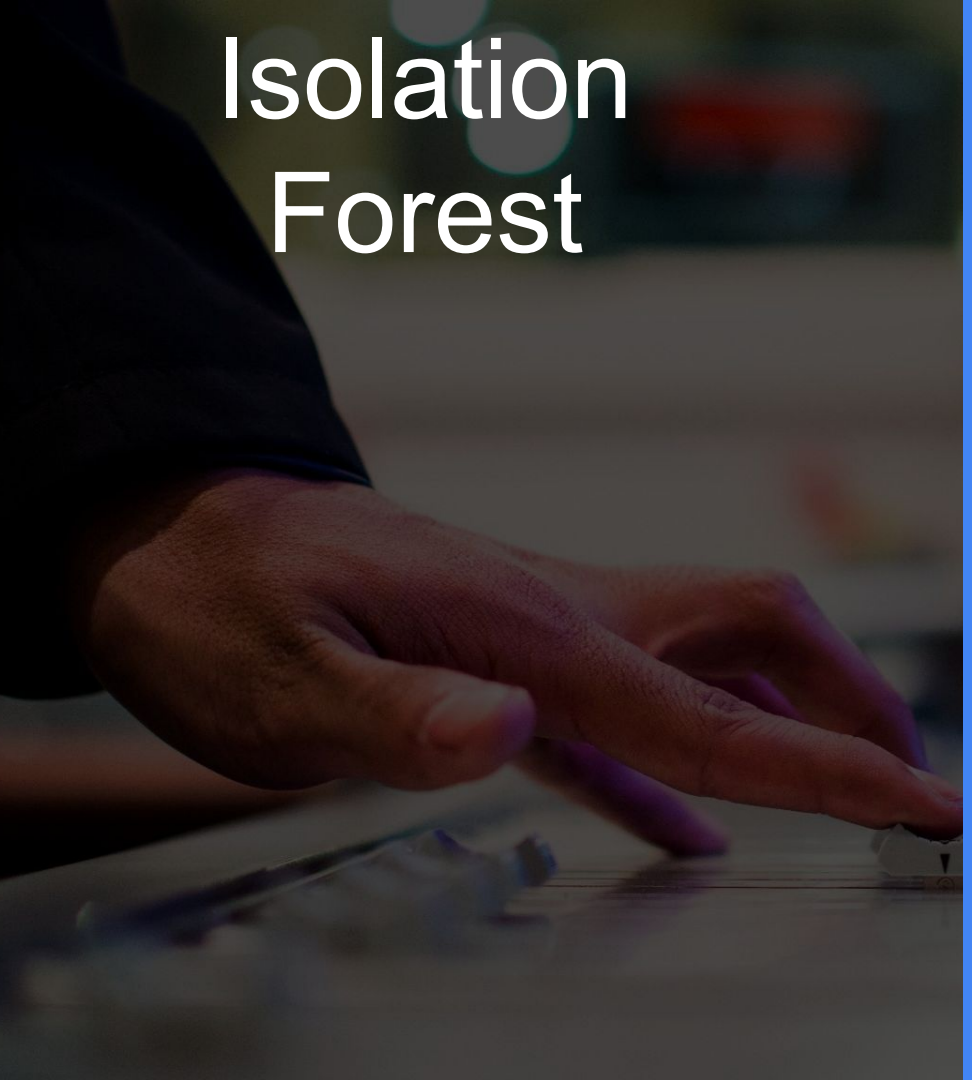
Implementation : Sequential Minimal Optimization (SMO)

Time :
Accuracy :

Reasons why Unsupervised is better ?

1. Imbalanced dataset :
2. Concept drift :

Isolation Forest

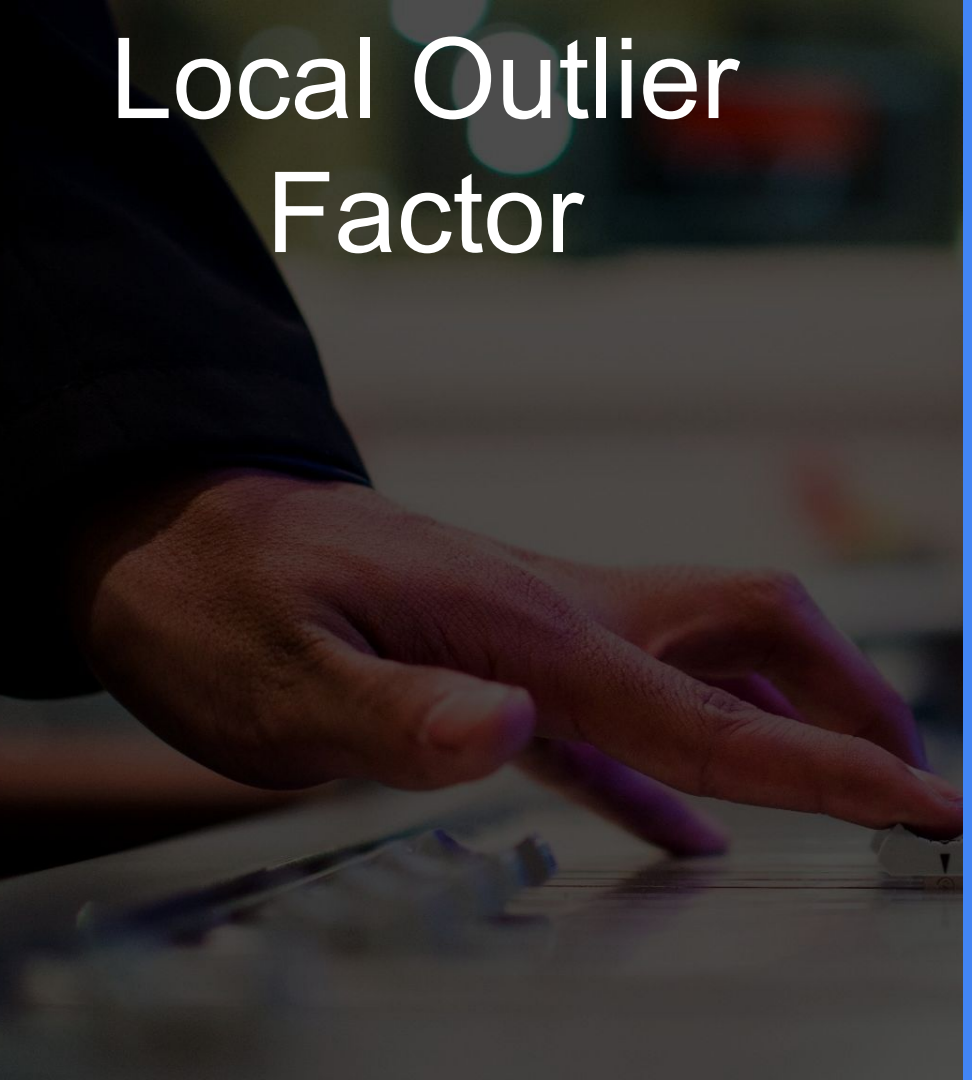


Isolation Forest:-

Time :

Accuracy :

Local Outlier Factor



Local Outlier Factor:-

Time :

Accuracy :

Previous Works / Dataset / References

1. Dataset : https://www.kaggle.com/mlg-ulb/creditcardfraud
2. Local Outlier Factor, Isolation Forest : https://ieeexplore.ieee.org/document/8741421
3. Logistic regression, SVM, Random Forest : https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8757212
4. https://en.wikipedia.org/wiki/Local_outlier_factor
5. https://content.linkedin.com/content/dam/engineering/site-assets/images/blog/posts/2019/08/IsolationForest1.png


Thank You..!