

Credit Card Fraud Detection



Presented by
Artem Ramus

Introduction

This data set is about credit card fraud detection. It is very likely that a bank will not share its client information with a data scientist. So, the feature names won't be available. It has a binary column that indicates if a transaction is fraudulent or not. This data set can be used for classification models.

Link to the dataset: https://github.com/rashida048/Datasets/blob/master/fraud_data.csv

Methodology

- The data set is checked for duplicates and null values
- Outliers and skewness were checked
- Constant, quasi-constant and highly-correlated features were deleted
- Cross-validation was used
- Oversampling technique SMOT was employed to compensate the imbalanced data set
- Hyper-parameters optimisation technique grid search was used

Performance of the Model

98% fraudulent credit card transactions were detected with random forest classifier using SMOTE technique to treat imbalanced data.

The end

Thank you for your attention!