10Alytics
data & strategy

Online Payments Fraud Detection Dataset Case Study

10Alytics
data & strategy

Case Study

Blossom Bank also known as BB PLC is a multinational financial services group, that offers retail and investment banking, pension management, asset management and payments services, headquartered in London, UK.

Problem

Blossom Bank wants to build a Machine Learning model to predict online payment fraud.

10Alytics
data & strategy

Data Dictionary

The below column reference:

- **step**: represents a unit of time where 1 step equals 1 hour
- **type**: type of online transaction
- **amount**: the amount of the transaction
- **nameOrig**: customer starting the transaction
- **oldbalanceOrg**: balance before the transaction
- **newbalanceOrig**: balance after the transaction
- **nameDest**: recipient of the transaction
- **oldbalanceDest**: initial balance of recipient before the transaction
- **newbalanceDest**: the new balance of the recipient after the transaction
- **isFraud**: fraud transaction

10Alytics
data & strategy

Data Science & Machine Learning Capstone Project

1. Problem definition: clearly articulate the problem that is to be solved with your data mining. How will the business benefit from your solution?

2. Perform exploratory data analysis in python.
    a) Visualize relationships between the label and some key features
    b) Explore correlations
    c) Conduct univariate and multivariate analysis as much as is feasible

3. Perform feature engineering
    a) Encoding categorical variables
    b) Create new features from existing features where necessary, depending on insights from your EDA

4. Model selection, training, and validation
    a) Train and test at least 2 supervised learning model

5. Model evaluation
    a) Analyse the results of your trained model
    b) What metrics are most important for the problem? For instance, should the business be more concerned with better results on false negatives or true positives?

6. Submission
    a) Publish your Jupyter notebook to your Github profile
    b) In the readme file, include a description of the project and summarize the steps you took and challenges you faced.
    c) share the link with your instructor

10Alytics
data & strategy

Perform your ML task and upload your code into your Github Repo.

Share link with @10Alytics and be prepared for a 10mins presentation