GarantiBBVA - Case

Introduction

The detection of fraud is one of the most recurrent problems in a large part of the current sectors. Many of the big companies lose large amounts of money due to fraud actions committed by their customers. In the banking context, this problem is especially pressing due to the difficulty of detecting these problems in time and acting accordingly. Furthermore, these fraud operations are easier to commit due to the increasing use of new technologies in the field of banking.

In this **"challenge"**, it is proposed to solve the classic problem of **fraud detection** in mobile transactions with *machine learning* techniques.

Dataset

- ***Training file (Train):*** It contains a subset of dataset transactions. This subset also contains the variable *isFraud*, which is the target variable. This dataset should be used to build a *machine learning* model capable of predicting whether a transaction is fraudulent or not.

Data structure

The data set consists of a table with numerous transactions. The meaning of each column is provided below:

- **step** *(int)* - maps a unit of time in the real world. In this case 1 step is 1 hour of time
- **type** *(text)* - CASH-IN, CASH-OUT, DEBIT, PAYMENT and TRANSFER.
    - CASH-IN: the process of increasing the balance of account by paying in cash to a merchant
    - CASH-OUT: it means to withdraw cash from a merchant which decreases the balance of the account
    - DEBIT: the process of sending the money from the mobile money service to a bank account
    - PAYMENT: process of paying for goods or services to merchants which decreases the balance of the account and increases the balance of the receiver

    - TRANSFER: process of sending money to another user of the service through the mobile money platform

- **amount** *(double)* - amount of the transaction in local currency.
- **nameOrig** *(text)* - customer who started the transaction
- **oldbalanceOrg** *(double)* - initial balance before the transaction
- **newbalanceOrig** *(double)* - new balance after the transaction
- **nameDest** *(text)* - customer who is the recipient of the transaction
- **oldbalanceDest** *(double)* - initial balance recipient before the transaction. Note that there is not information for customers that start with M (Merchants).
- **newbalanceDest** *(double)* - new balance recipient after the transaction. Note that there is not information for customers that start with M (Merchants).
- **isFraud** *(int)* - if the transaction is fraudulent or not (0/1)
- **isFlaggedFraud** *(int)* - an attempt to transfer more than 200.000 in a single transaction (0/1)