**Financial Fraud Detection**

Technology is growing exponentially everyday, and this advancement is affecting each and every major business sector. It is very evident that this is the digital age and business interactions have completely changed.

The increasing demand for digital banking experience by the customers has completely changed the way in which the banking sector operates. **Mobile Banking** has its hand in almost every aspect of the banking industry.

Most of our transactions nowadays, be it account transfer, making deposits, monitoring our spendings and earnings are done through Mobile Banking. All your transactions are just a click away, which seemed very unlikely a few years ago.

It seems everything’s great, is it ?

One of the major concerns that consumers have with digital banking is security. The fear of data breach or any kind of fraud increases the need for services that keep user’s data secure and be fraud proof.The financial services industry and the industries that involve financial transactions are suffering from fraud-related losses and damages.

**So, the challenge for industry players is to implement real-time claim assessment and improve the accuracy of fraud detection.**

**Problem Statement:**

**Can we predict whether there has been a fraudulent transaction using Machine Learning techniques?**

**Steps:**

1. **Exploratory Data Analysis**
2. **Data Cleaning**
3. **Feature Engineering**
4. **Data Visualization**
5. **Machine Learning Approach to detect fraud**
6. **Model Improvement**

**Data :**

Transactional data containing more than 6 million rows has been provided. All the transactions have been done through mobile banking. It contains 11 columns, and the description of each of the column has been provided below:

This is a sample of 1 row with headers explanation:

1,PAYMENT,1060.31,C429214117,1089.0,28.69,M1591654462,0.0,0.0,0,0

* **step** - maps a unit of time in the real world. In this case 1 step is 1 hour of time. Total steps 744 (30 days simulation).
* **type** - CASH-IN, CASH-OUT, DEBIT, PAYMENT and TRANSFER.
* **amount** - amount of the transaction in local currency.
* **nameOrig** - customer who started the transaction
* **oldbalanceOrg** - initial balance before the transaction
* **newbalanceOrig** - new balance after the transaction
* **nameDest** - customer who is the recipient of the transaction
* **oldbalanceDest** - initial balance recipient before the transaction. Note that there is not information for customers that start with M (Merchants).
* **newbalanceDest** - new balance recipient after the transaction. Note that there is not information for customers that start with M (Merchants).
* **isFraud** - This is the transactions made by the fraudulent agents inside the simulation. In this specific dataset the fraudulent behavior of the agents aims to profit by taking control or customers accounts and try to empty the funds by transferring to another account and then cashing out of the system.
* **isFlaggedFraud** - The business model aims to control massive transfers from one account to another and flags illegal attempts. An illegal attempt in this dataset is an attempt to transfer more than 200.000 in a single transaction.