FROSTHACK

ACE DYNAMICS

Credit Card fraud Prediction

-Dev Haral

IIT Mandi

Ace Dyanamic Explained

- What is this Technology?
- What can we expect from AD?
- Why we need this?
- What are the benefits of AD?
- Technical breakdown of AD.

Technical Breakdown!

1

ML Model

Using Logistic regression and Kaggle data set buid model.

2

Predict

Kaggle Notebook used to buid and predict test data.

3

Deploy

Using flask/html/css deploy model .

4

Prection

Giving 30 input as transaction model will predict data.

5

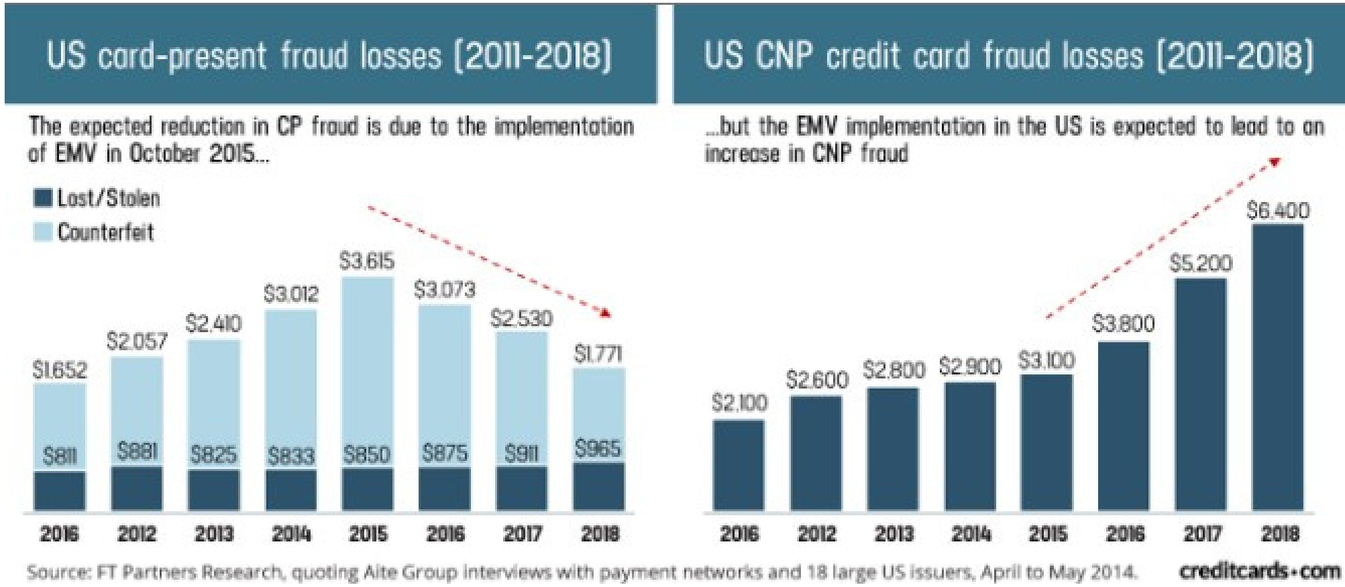
Website

A stylish website to present entire Project.

Introduction

In today's world, we are on the express train to a cashless society. According to the World Payments Report, in 2016 total non-cash transactions increased by 10.1% from 2015 for a total of 482.6 billion transactions! That's huge!

This is now becoming a serious problem since most of the time, a person who has become a victim of this fraud don't have any idea about what has happened until the very end. So in this project, what we have tried is to create a Web App for the detection of such type of frauds with the help of Machine Learning.



Model Training

1. Understanding the data and related constraints
2. Preprocessing data
3. Model Architecture
4. Post Training
5. Creating Web App

Understanding the data and related constraints

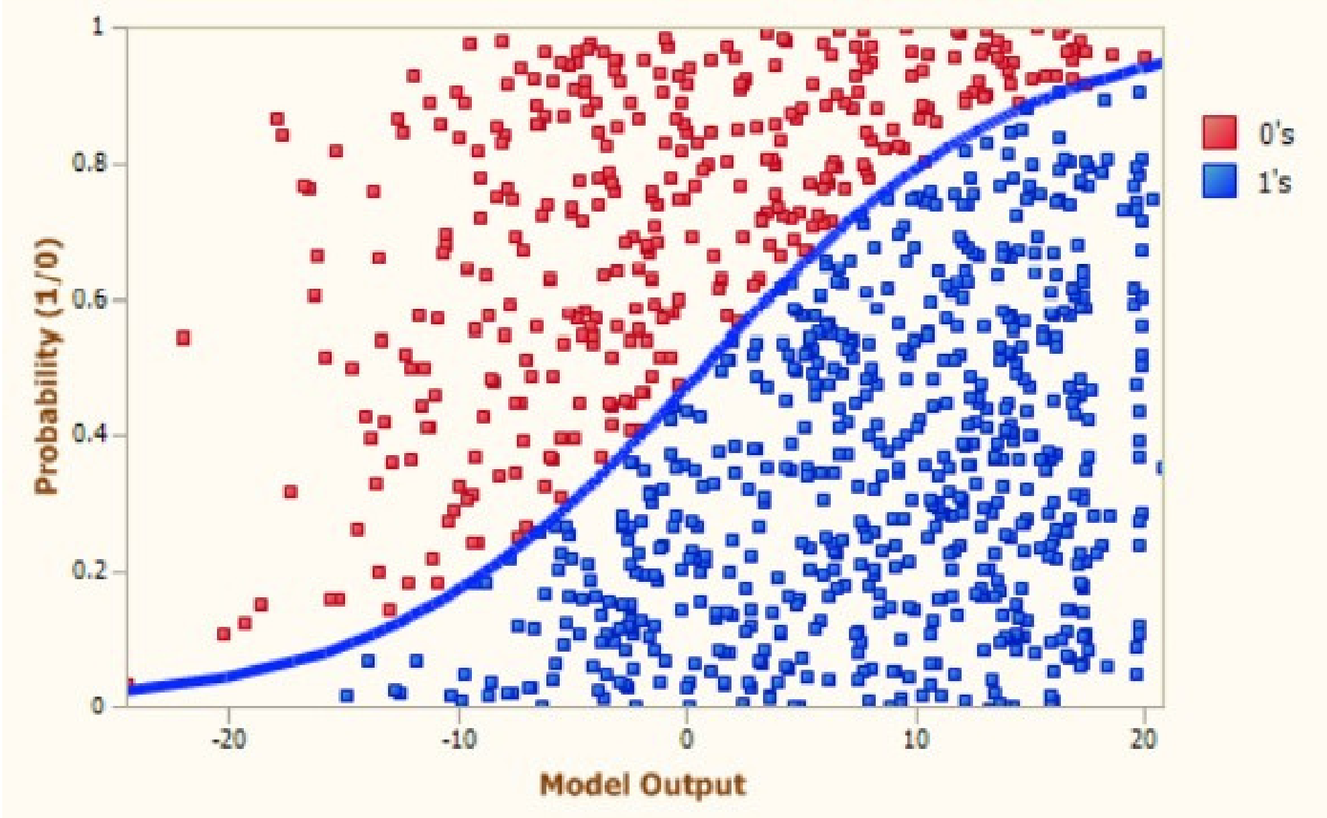
- Since the data for this project is very unbalanced due to the fact that number of cases of Fraud transactions are very low in comparison to number of cases of Valid transactions makes the model training a bit hectic.
- For this reason, we use a different type of metric which will give us much more important information about what our model has learned. Actually, what we do is we print the classification matrix for our model predictions and then we judge our model based on that matrix.
- Precision and Recall are two of the derivatives of a confusion matrix, we will consider them both though, but only Recall will come in handy for us since high Recall will ensure that no fraud value gets detected to be a valid one. Also, Precision do the vice-versa

PREPROCESSING DATA

- Balancing data Since the data is very imbalanced, we will be using some Undersampling and Oversampling techniques. As the name suggests, Undersampling is use to reduce the samples from majority class and Oversampling is used to increase the samples from minority class.
- Scaling features Now, even though almost all of the features are dimensionally reduced using some dimensionality reduction technique, two of the features are in their original form.
- Splitting the data Now, since we needed to save some entries from the data for our testing purpose, we will now be splitting the data into two parts, namely, train and test .

Model Architecture

We will be using a Logistic Regression classifier for our project. A logistic Regression model is used to predict the probability of a certain class or event existing. We then decide the class from which the entry belongs by using a threshold value. This threshold value is decided by manipulating the Precision-Recall tradeoff as explained above.



Creating Web App

Brief Introduction to a Web App

A web application is an application software that runs on a web server, unlike computer-based software programs that are stored locally on the Operating System of the device. Web applications are accessed by the user through a web browser with an active internet connection. These applications are programmed using a client–server modeled structure—the user is provided services through an off-site server that is hosted by a third-party. The third party whose services we are using is Heroku, Heroku is a great place to launch your apps upto a limited size. We could not do this project without the support of Heroku who gave us the opportunity to host our web app.

Conclusion

Credit Card is a great tool to pay money easily, but as with all the other monetary payment tools, reliability is a issue here too as it is subjected to breach and other frauds. To encounter this problem, a solution is needed to identify the patterns in the transactions and identify the ones which are fraud, so that finding such transactions beforehand in future will be very easy.
Machine Learning is a great tool to do this work since Machine Learning helps us in finding patterns in the data. Machine Learning can help producing great results if provided enough amount of data. Also, with further advances in the technology, Machine Learning too will advance with time, it will be easy for a person to predict if a transaction is fraud or not much more accurately with the advances.