Synopsis

Project Title

Predicting customer behaviour towards loyalty program

Capstone Project Team

Ankit Mathur

Apurv Salunke

Rinesh Rajan

Vardharajan G P

Raghul Sharan R

Capstone Project Mentor

Ankush Bansal

Date: 27th Sept 2020

Batch: DSE Feb 2020 Pune - Group 5

Group: Project Group 2

Problem Statement

A credit card company named XYZ regularly helps its merchants understand their data better by providing machine learning and analytics consulting. ABC is an established Brick & Mortar retailer that frequently conducts marketing campaigns for its diverse product range. As a merchant of XYZ, they have sought XYZ to assist them in their discount marketing process.

ABC's promotions are shared across various channels which includes email, notifications, etc. A number of these campaigns include coupon discounts that are offered for a specific product/range of products.

The retailer would like the ability to predict whether customers redeem the coupons received across channels, which will enable the retailer's marketing team to accurately design coupon construct, and develop more precise and targeted marketing strategies.

In this problem the data available contains the details of a sample of campaigns and coupons used in previous campaigns. Our task is to predict the probability for the next 10 campaigns in the test set given the information for the last 18 campaign for each coupon and customer combination.

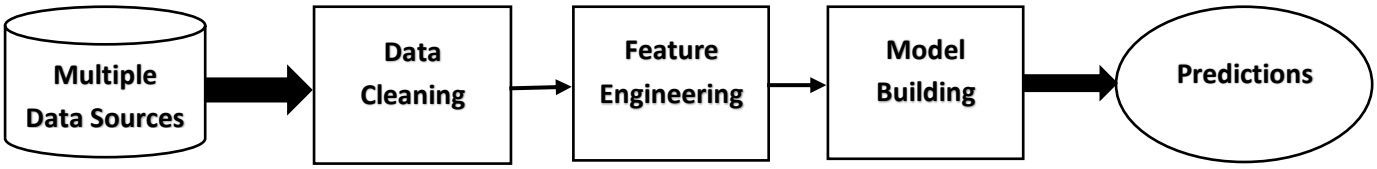
Data Description

The data available in this problem contains the following data tables mentioned below required for our analysis and prediction

- **train.csv:** Train data containing the coupons offered to the given customers under the 18 campaigns
- **campaign_data.csv:** Campaign information for each of the 28 campaigns
- **coupon_item_mapping.csv:** Mapping of coupon and items valid for discount under that coupon
- **customer_demographics.csv:** Customer demographic information for some customers
- **customer_transaction_data.csv:** Transaction data for all customers for duration of campaigns in the train data
- **item_data.csv:** Item information for each item sold by the retailer
- **test.csv:** Contains the coupon customer combination for which redemption status is to be predicted

Proposed Methodology

Below mentioned is the overview of our machine learning pipeline which consists of various steps to train our model.



Evaluation Metrics

Area under the ROC curve between the predicted probability and the observed target would be used as the evaluation metric for our classification model.