



# Credit Card Lead Prediction (Approach Document)

30.05.2021

---

Siddharth Ranjan

Mu Sigma - 3years exp as Decision Scientist

Bangalore

9776020503

## Overview

Happy Customer Bank is a mid-sized private bank that deals in all kinds of banking products, like Savings accounts, Current accounts, investment products, credit products, among other offerings.

The bank also cross-sells products to its existing customers and to do so they use different kinds of communication like tele-calling, e-mails, recommendations on net banking, mobile banking, etc.

**Current problem:** the Happy Customer Bank wants to cross sell its credit cards to its existing customers. The bank has identified a set of customers that are eligible for taking these credit cards.

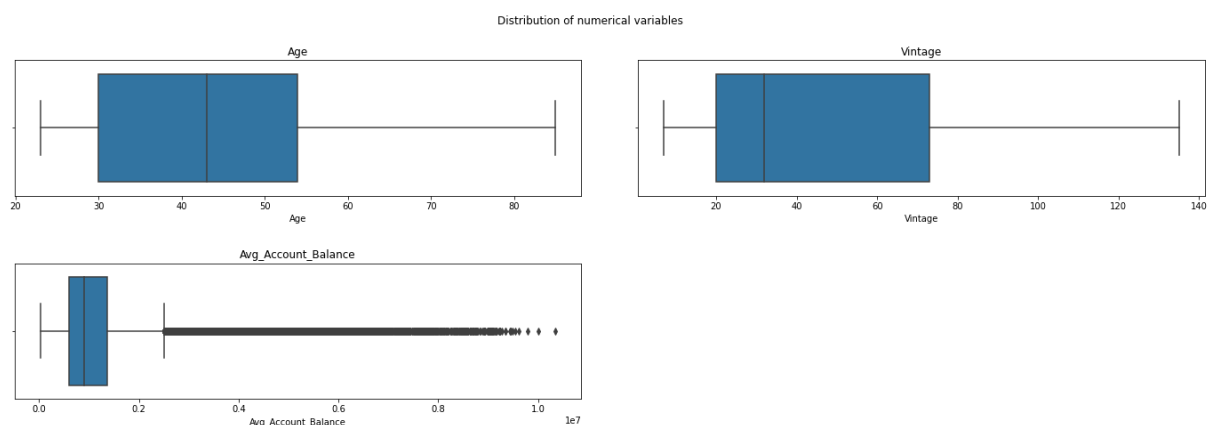
## Approach

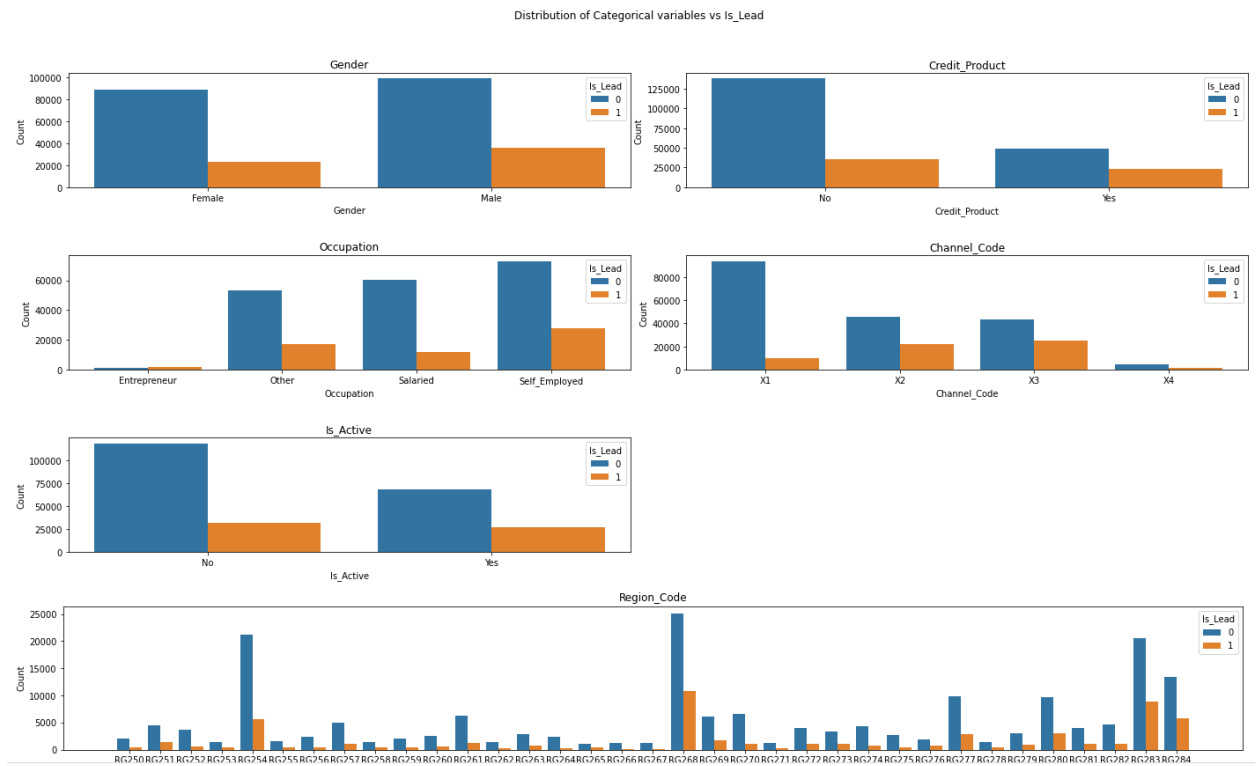
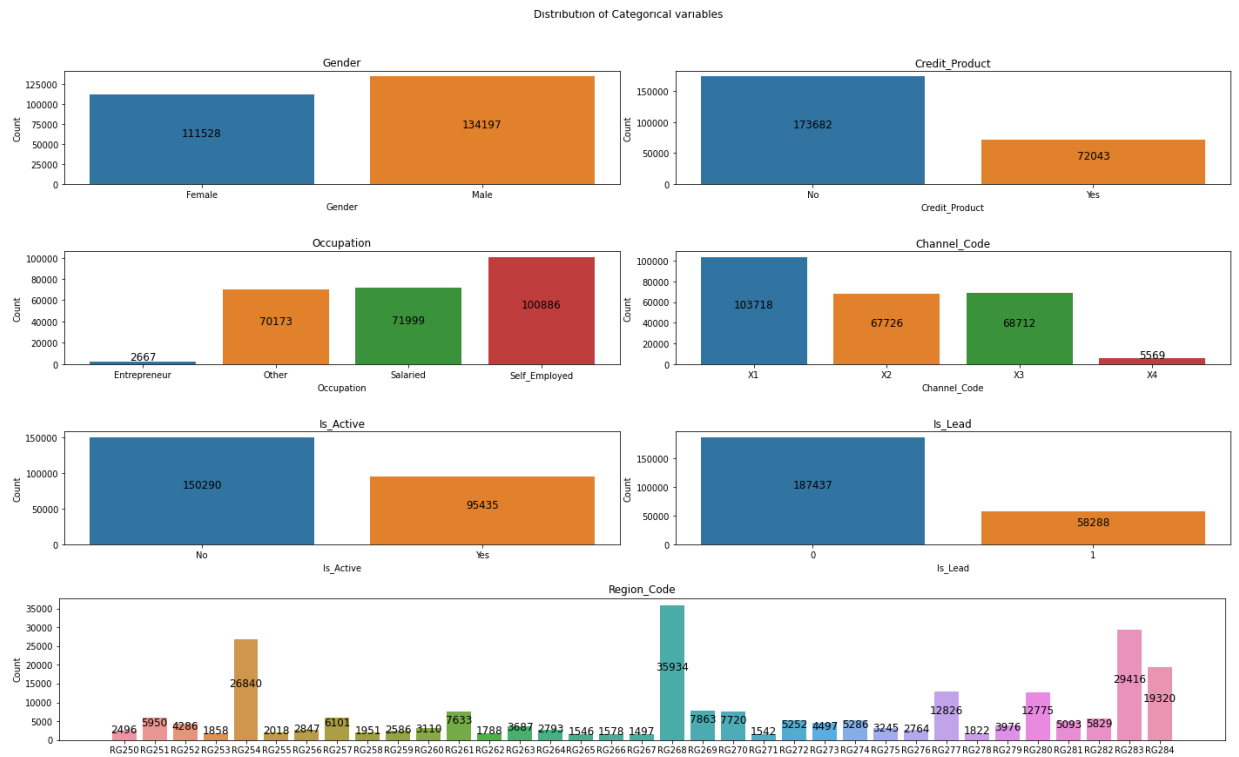
### I. Data Wrangling

To get the information on how the data looks like and check and treat missing values in the data. Current data had only one variable having missing value which is Credit\_Product

### II. Exploratory Data Analysis

EDA was done to see the distribution and trend of the data. Please find below some screenshot





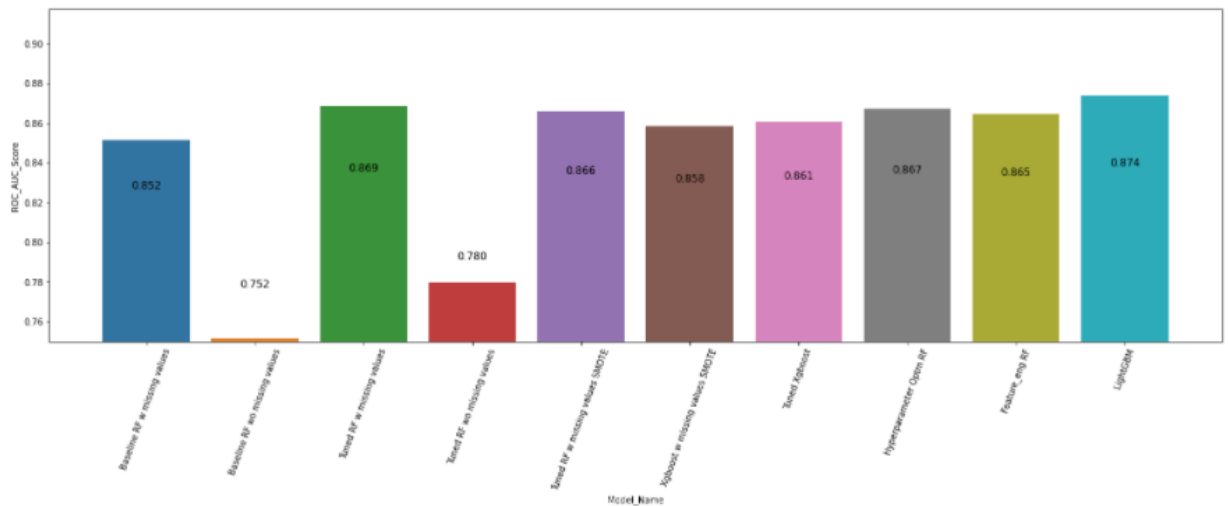
This could be further drilled down for important insights but was avoided, since the only goal was prediction.

### III. Model Training

In total we trained 9 models. Details are given below:

	Description	Model used
1	Default parameters and no replacement in missing value	Random Forest
2	Default parameters and replacement in missing value	Random Forest
3	Tuned parameter and no replacement in missing value	Random Forest
4	Tuned parameters and replacement in missing value	Random Forest
5	Tuned parameter, handling imbalance with SMOTE and no replacement in missing value	Random Forest
6	Tuned parameters and replacement in missing value	XGboost
7	Tuned parameter, handling imbalance with SMOTE and no replacement in missing value	XGboost
8	Hyperparameter optimization using Random search	Random Forest
9	Feature engineering, tuned parameter and no replacement in missing value	Random Forest
10	Tuned parameter and no replacement in missing value, Kfold validation	Light GBM model
11	Tuned parameter and no replacement in missing value, Kfold validation	XGboost
12	Ensemble of above two models	Light GBM model + XGboost

## Snapshot of Models with Scores



The one with the highest score is selected as champion model and 2nd highest as challenger. Both result is cross checked in the solution checker and selected accordingly