11.3 Final Project Step 3

Sharath Chandra Tummanapally

11/17/2021

Introduction

This project is an attempt to forecast how we can decrease loan defaults of Personal Loans. I built a model for this loan prediction. This problem would be interesting to solve because it's existing problem in loan business and eventually by predicting the defaulter, it could be beneficial to the lending organization. This prediction might help any bank to be prepared for any foreseen pandemics in future.

The problem statement you addressed

One of the most difficult challenges is to reduce the number of loan defaulters in a personal lending industry. Personal loan defaults has been a significant problem especially during this COVID-19 period in this industry. If customers pay their interest on time and do not default on their loans, the personal Lending firms will be lucrative. If more loans default, the company's profitability will suffer. As a result, it is very helpful that the lender be able to forecast the possibility of a borrower defaulting before offering loan to the customer.

How you addressed this problem statement

There are three different datasets used:

1. Customer Loan Details over years from a bank.
2. Unemployment rate over the years in India.
3. Lending Interest rate of Indian banks over the years.

All the three data sets are combined into a single data set that has all the available information for a model to predict the outcome. I used logistic regression model to identify the relation between variable Risk_Flag based on the variables Income, Interest Rate, Unemployment Rate and Age during the loan defaulted year

CustLoanModel <- glm(Risk_Flag ~ Income + InterestRate + UnempRate + Age , data = CustLoanInfo_df, family = binomial())

Analysis

As per my analysis, when the model's summary was applied, we can see that the bank's location's unemployment rate and the loan's interest rate have biggest impact on loan to default.

Implications

According to the datasets I utilized for my analysis, unemployment rate and interest rate are two important characteristics that influence the likelihood of loans defaulting in a personal lending bank. A higher interest rate means higher monthly payments that are difficult to pay back. When a bank issues a personal loan, it must consider the customer's job situation as well as various other financial facts in order to authorize the loan amount.

Limitations

There could be various relief programs that are offered to customers to help ease loan payments for certain period of time. In this model and project, we didn't consider those programs.

Concluding Remarks

The model predicts that the Unemployment rate for a location and higher interest rate can decides the outcome of the loan whether its going to be defaulted. Also the accuracy of the model is more than 80% which implies that it is not a bad model with the data that we have at hand.