Executive Summary Report
Loan Approval Project

By
HRISAV BHOWMICK
PGPDS
PRAXIS BUSINESS SCHOOL
KOLKATA, INDIA

Submitted on
February 2021

Loans are the core business of banks. The major profit comes directly from the loan's interest. The banks grant a loan after an intensive process of verification and validation. However, they still don't have the assurance whether the applicant will be able to repay the loan with no difficulties.

To explore the problem we have used the Analytics Vidhya dataset. Dream Housing Finance company deals in all home loans. They have presence across all urban, semi urban and rural areas. Initially the Customers apply for a home loan and then the company validates the customer if he/she is eligible for loan.

Company wants to automate the loan eligibility process (real time) based on customer details provided while filling an online application form. These details are Gender, Marital Status, Education, Number of Dependents, Income, Loan Amount, Credit History and others. To automate this process, they have given a problem to identify the customer segments that are eligible for loan, so that they can specifically target these customers.

To solve the issue, firstly, we try to handle the missing values in the dataset by filling with mean, median or mode values based on the requirements. Secondly, we try to perform exploratory data analysis to understand the distribution of the data. We try to understand things like whether loan is majorly approved in case of male or female, or in case of a graduate or non-graduate, or based on the property area. Also we try to find out if the loan amount being provided is higher in case of graduates or in case of married applicants. Thirdly, we add some new features and then try to take the relevant features required for model building, and work on them. If any feature is having categorical or string type values, we transform them into numerical type so that it can be utilised for model building.

After all the pre-processing steps are done, finally, we build a predictive model to predict if an applicant will be able to repay the lending company or not. In this project we have built three models, and we choose the one which gives better performance results according to the business requirement.

Some of the key findings are:

- Median Loan Amount provided to the Rural Area is slightly higher than Urban. But there are some people in Urban who have higher loan amounts compared to Rural or Semi urban.

- Graduates, Married people have a higher chance of getting loan approved.

- When the number of dependents is 2 there is more chance of getting loan approval compared to when the number of dependents is 1.

- There is a very rare chance of getting a loan if the person has no credit history.

- Most of the people are from the Semiurban area, and highest Loan acceptance is from that area.