



LOAN PREDICTION BASED ON CUSTOMER BEHAVIOR

Manoj Boyapati

Introduction

- In India, the number of people applying for the loans gets increased for various reasons in recent years. The bank employees are not able to analyze or predict whether the customer can payback the amount or not (good customer or bad customer) for the given interest rate.
- This project involves predicting the customers eligibility for the loan. The model is prepared using various columns in the dataset like age, income, experience, risk flag, etc.
- The dataset has been acquired from Kaggle. This dataset is owned by “Univ.AI” and was created for the hackathon organized by them. All the values in the dataset has been provided by the applicants at the time of loan application.

DETAILS OF ALL THE COLUMNS IN THE DATASET

Column	Description
income	Income of the user
age	Age of the user
experience	Professional experience of the user in years
profession	Profession
married	Whether married or single
house_ownership	Owned or rented or neither
car_ownership	Does the person own a car
risk_flag	Defaulted on a loan
current <i>job</i> years	Years of experience in the current job
current <i>house</i> years	Number of years in the current residence
city	City of residence
state	State of residence

DOMAIN AND SOLUTION

- The main target of the project is to find if the customer is eligible for the loan or not. Logistic regression is used to create the model. Risk flag is one of the most important values to determine the loan eligibility. Risk flag says if the customer has defaulted on a loan previously. There is a greater chance of the customer defaulting the loan if he has done it previously.
- Data visualizations on the dataset will also be done like marital status ratio, frequency of each profession, average age of profession, etc.