

LENDING CLUB

CASE STUDY

- By Deepak Kumar Jha

Case study objectives

- Identification of Loan applicant traits that tend to default paying back
- Understand the driving factors behind loan default
- Loan lending organizations want to utilize this knowledge to assess the risk of new loan applicants

Methodology used

- Understanding Dataset
- Data Cleaning
- Data Preparation
- Exploratory Data Analysis

Data Understanding

- Most of variables are not associated in identifying the default as they come in picture when the loan is approved.
- we are focused if we want to approve loan in starting or not. We can categorize all remaining variables in 3 category
 - Customer Demographics variables
 - Loan and Investor variables
 - Customer Behaviour variables
- To meet business objective, we will only consider the Customer Demographic and Loan variables and we will ignore the Customer behaviour variables as these will not be known during the time of application.

Data Cleaning

- There are 54 columns have all missing values. We removed it.
- Removed customer behavior variables
- Removed few unused customer demographics variables
- Removed columns having single valued
- Filled up missing values
- Standardize few columns data such as emp_length, subgrade, int_rate etc.

Data Preparation

- Checked/Removed outliers
- Derived multiple columns
 - annual_inc_groups
 - int_rate_groups
 - loan_amnt_groups
 - issue_year ..etc.

Exploratory Data Analysis

There is a more probability of defaulting when:

- Applicants taken a loan for small business and the loan amount is greater than 14k
- Applicants taken a loan in the range 30k - 35k and charged interest rate between 15-17 %
- Interest rate for charged off loans is pretty high than that of fully paid loans in all the loan_amount groups
- Applicants whose home ownership is MORTGAGE and taken a loan of 14-16k
- Loan issue month is December and loan amount is between 12k - 13.5k
- Grade is F and loan amount is between 17k-20k
- Employment length is 10yrs and loan amount is 12k-14k
- Loan is verified and loan amount is above 16k
- Applicants taking loan for home improvement and have income of 60k-65k
- Applicants whose home ownership is MORTGAGE and have income of 63-68k
- Applicants annual income is between 112k-140k and loan amount is between 16k-20k
- Applicants who receive interest at the rate of 21-24% and have an income of 65k-72k
- Grade is G and interest rate above 20%

Final Results

- Low grade loans have high tendency to default. Grading system is working as expected.
- Loans having higher interest rate have more defaulters. Check the background of applicant if interest rate is high.
- When purpose of loan is debt consolidation check applicant thoroughly as it has high tendency to default.