

CASE STUDY - BANKING
Credit Risk Analytics

OVERVIEW OF CREDIT SCORING:

Credit scoring is perhaps one of the most "classic" applications for predictive modeling, to predict whether or not credit extended to an applicant will likely result in profit or losses for the lending institution. There are many variations and complexities regarding how exactly credit is extended to individuals, businesses, and other organizations for various purposes (purchasing equipment, real estate, consumer items, and so on), and using various methods of credit (credit card, loan, delayed payment plan). But in all cases, a lender provides money to an individual or institution, and expects to be paid back in time with interest commensurate with the risk of default.

Credit scoring is the set of decision models and their underlying techniques that aid lenders in the granting of consumer credit. These techniques determine who will get credit, how much credit they should get, and what operational strategies will enhance the profitability of the borrowers to the lenders. Further, they help to assess the risk in lending. Credit scoring is a dependable assessment of a person's credit worthiness since it is based on actual data.

A lender commonly makes two types of decisions: first, whether to grant credit to a new applicant, and second, how to deal with existing applicants, including whether to increase their credit limits. In both cases, whatever the techniques used, it is critical that there is a large sample of previous customers with their application details, behavioral patterns, and subsequent credit history available. Most of the techniques use this sample to identify the connection between the characteristics of the consumers (annual income, age, number of years in employment with their current employer, etc.) and their subsequent history.

Typical application areas in the consumer market include: credit cards, auto loans, home mortgages, home equity loans, mail catalog orders, and a wide variety of personal loan products.

MODEL BUILDING:

When the training data set on which the modeling is based contains a binary indicator variable of "Paid back" vs. "Default", or "Good Credit" vs. "Bad Credit", then Logistic Regression models are well suited for subsequent predictive modeling. Logistic regression yields prediction probabilities for whether or not a particular outcome (e.g., Bad Credit) will occur. Furthermore, logistic regression models are linear models, in that the logit-transformed prediction probability is a linear function of the predictor variable values. Thus, a final score card model derived in this manner has the desirable quality that

the final credit score (credit risk) is a linear function of the predictors, and with some additional transformations applied to the model parameter, a simple linear function of scores that can be associated with each predictor class value after coarse coding. So the final credit score is then a simple sum of individual score values that can be taken from the scorecard

BUSINESS OBJECTIVES:

The application of scoring models in today's business environment covers a wide range of objectives. The original task of estimating the risk of default has been augmented by credit scoring models to include other aspects of credit risk management: at the pre-application stage (identification of potential applicants), at the application stage (identification of acceptable applicants), and at the performance stage (identification of possible behavior of current customers). Scoring models with different objectives have been developed. They can be generalized into four categories as listed below.

1. MARKETING ASPECT:

Purposes:

Identify credit-worthy customers most likely to respond to promotional activity in order to reduce the cost of customer acquisition and minimize customer dissatisfaction.

Predict the likelihood of losing valuable customers and enable organizations to formulate effective customer retention strategy.

Examples:

Response scoring: The scoring models that estimate how likely a consumer would respond to a direct mailing of a new product.

Retention/attrition scoring: The scoring models that predict how likely a consumer would keep using the product or change to another lender after the introductory offer period is over.

2. APPLICATION ASPECT

Purposes

Decide whether to extend credit, and how much credit to extend.

Forecast the future behavior of a new credit applicant by predicting loan-default chances or poor repayment behaviors at the time the credit is granted.

Examples:

Applicant scoring: The scoring models that estimate how likely a new applicant of credit will become default.

3. PERFORMANCE ASPECT

Purpose:

Predict the future payment behavior of existing debtors in order to identify/isolate bad customers to direct more attention and assistance to them, thereby reducing the likelihood that these debtors will later become a problem.

Examples:

Behavioral scoring. Scoring models that evaluate the risk levels of existing debtors.

4. BAD DEBT MANAGEMENT

Purpose:

Select optimal collections policies in order to minimize the cost of administering collections or maximizing the amount recovered from a delinquent's account.

Examples:

Scoring models for collection decisions: Scoring models that determine when actions should be taken on the accounts of delinquents and which of several alternative collection techniques might be more appropriate and successful.

Thus, the overall objective of credit scoring is not only to determine whether the applicant is credit worthy, but also to attract quality credit applicants who can subsequently be retained and controlled while maintaining an overall profitable portfolio.

DATA AVAILABLE:

➤ Bankloans.csv

The data contains the credit details about credit borrowers:

Data Description:

age - Age of Customer

ed - Education level of customer

employ: Tenure with current employer (in years)

address: Number of years in same address

income: Customer Income

debtinc: Debt to income ratio

creddebt: Credit to Debt ratio

othdebt: Other debts

default: Customer defaulted in the past (1= defaulted, 0=Never defaulted)

Project Approach:

Table shows the list of data pre-processing, analysis, visualization and model building techniques that needs to be implemented to complete the project

Task	Task Details	Analytical Techniques	Visualization Techniques
Data Manipulation & Preparation	1. Perform required data manipulation and cleaning.	1. Descriptive statistics and outlier analysis	1. Bar Plot, Seaborn (Heatmap), boxplots
	2. Perform Uni variate and Bi variate analysis	2. Independent t-test and VIF check to get important features	2. Boxplot segmentation
Model Building & Performance Check	Create Model and Asses the performance of the models	Build Logistic and Decision tree models, used hyper parameter to fine tune the model.	Visualize the decision tree using graphviz package

Submission Guidelines: Best Practice & Procedures

- The python file should be submitted in IPYNB format
- A PPT presentation outlining your findings must be included in the submission
- To submit the files, create a folder name with your email ID
- Finally upload the folder containing your files to the designated Batch folder on G-Drive
- For submission kindly watch the attached recording