

Credit Risk Modeling

This competition concerns loan data. When a customer applies for a loan, banks and other credit providers use statistical models to determine whether or not to grant the loan based on the likelihood of the loan being repaid. The factors involved in determining this likelihood are complex, and extensive statistical analysis and modeling are required to predict the outcome for each individual case. You must implement a model that predicts loan repayment or default based on the data provided.

The dataset used in this competition consists of synthetic data that was generated specifically for use in this project. The data is designed to exhibit similar characteristics to genuine loan data.

Description


In this competition, you must explore and cleanse a dataset consisting of over 111,000 loan records to determine the best way to predict whether a loan applicant will fully repay or default on a loan. You must then build a machine learning model that returns the unique loan ID and a loan status label that indicates whether the loan will be fully paid or charged off.

Dimensionless Technologies Private Limited

B/303, Silicon Valley, Balewadi, Pune – 411045

Visit us at: www.dimensionless.in

 info@dimensionless.in

 9923170071, 8108094992

CIN - U74900PN2014PTC15148



Data files

You can download the training data from [here](#).

The dataset consists of the following fields:


- Loan ID: A unique Identifier for the loan information.
- Customer ID: A unique identifier for the customer. Customers may have more than one loan.
- Loan Status: A categorical variable indicating if the loan was paid back or defaulted.
- Current Loan Amount: This is the loan amount that was either completely paid off, or the amount that was defaulted.
- Term: A categorical variable indicating if it is a short term or long term loan.
- Credit Score: A value between 0 and 800 indicating the riskiness of the borrower's credit history.
- Years in current job: A categorical variable indicating how many years the customer has been in their current job.
- Home Ownership: Categorical variable indicating home ownership. Values are "Rent", "Home Mortgage", and "Own". If the value is OWN, then the customer is a home owner with no mortgage
- Annual Income: The customer's annual income
- Purpose: A description of the purpose of the loan.

Dimensionless Technologies Private Limited

B/303. Silicon Valley, Balewadi, Pune – 411045

Visit us at: www.dimensionless.in

 info@dimensionless.in

 9923170071, 8108094992

CIN - U74900PN2014PTC15148



- Monthly Debt: The customer's monthly payment for their existing loans
- Years of Credit History: The years since the first entry in the customer's credit history •

Months since last delinquent: Months since the last loan delinquent payment

- Number of Open Accounts: The total number of open credit cards
- Number of Credit Problems: The number of credit problems in the customer records.
- Current Credit Balance: The current total debt for the customer
- Maximum Open Credit: The maximum credit limit for all credit sources.
- Bankruptcies: The number of bankruptcies
- Tax Liens: The number of tax liens.

Evaluation criteria

To achieve a passing grade, the accuracy of the model has to be at least 70 (percent).