Problem Statement

Bharat Bank is a banking institution. Every year, a large number of people take home loans from Bharat Bank, and many of them default on their loans which results in huge losses for the bank. Given below are the details of some loan applications and whether the borrower defaulted or not.

Based on the given dataset, you have to predict whether the applicant will default or not.

The training dataset consists of the following fields:

- LOAN_ID: Loan ID
- DEFAULTER: 1 = applicant defaulted on loan or seriously delinquent;
                        0 = applicant repaid the loan fully
- AMOUNT: Amount of the loan request
- DUE_MORTGAGE: Amount due on existing mortgage
- VALUE: Value of property under consideration
- REASON: Reason for loan request
- OCC: Occupational categories
- TJOB: Years spent at present job
- CONVICTED: Has the applicant ever been criminally convicted.

- DCL: Number of delinquent credit lines
- CLT: Age of first credit line in months
- VAR_1, VAR_2, VAR_3: Anonymized variables
- CL_COUNT: Total number of credit lines
- RATIO: Loan Amount to income ratio

The evaluation metric 'auc-roc' will be used for evaluation.

The test-train split is 40:60.