PREDICTION MODEL FOR LOAN DEFAULT ASSESSMENT USING MACHINE LEARNING TECHNIQUES

Submitted by

1. Bhuvaneshwaran Ravi (002275116)
2. Serlin Tamilselvam (002267722)

- Our prediction model is developed using Gradient Boosting classifier. The conceptual idea behind this classifier is to pick an algorithm and make tweaks to it with various regularization schemes, this process improves the learning ability of the model in a gradual and additive fashion. This classifier is particularly effective at classifying complex dataset such as Banking and Financing.

- The results predicted by our model on the probability a customers will default on their loan repayment is accurate ~76% of the time (this is validated against a pre-defined validation dataset fed into the model during the train/test phase).

- Our model identified a potential set of customers who maintains a significant account balance and has regular intervals of credit transactions on their account, has high probability of loan repayment. The customers with less account transactions and maintains a very low account balance for a significant time before the loan request date, has a high probability of loan defaulting. The model is also capable of adjusting the prediction parameters of its features based on the future dataset fed into it.

- In total we extracted 21 features from the provided dataset which yields maximum accuracy from our model, but we were able to identify certain features to be most crucial to predict the customer's probability of loan defaulting,

1. Total number of credit transactions
2. Average amount credited to the account
3. Total number of debit transactions
4. Average amount debited from the account
5. Requested loan amount
6. Total amount of all credit transactions
7. Days between loan request and account transactions
8. Current balance in the customer's account
9. Loan request date
10. Total amount of all debit transactions

- The idea behind choosing these particular features is to train our prediction model on the bank's existing customers transactional behavior and their loan outcomes, So that any potential loan seekers can be evaluated and risk-assessed primarily based on their transactional history.