IBM PROJECT

SMART LENDER – APPLICANT CREDIBILITY PREDICTION FOR LOAN APPROVAL

TEAM ID - PNT2022TMID35325

Siva Deepak S P

Sandeep Saghar R

Kesavan P

Sundarrajan P

ALGORITHM FOR THE LOAN CREDIBILITY PREDICTION SYSTEM

AUTHOR

Soni P M, Varghese Paul

METHODOLOGY

Classification is the most suitable predictive modelling technique in data mining to predict the loan repayment capability of a customer in a banking industry. There are various methods to improve the accuracy of a classification algorithm. The accuracy of random forest classification algorithm can be improved using Ensemble methods, Optimization techniques and Feature selection. Various feature selection methods are available. In this research work a novel hybrid feature selection algorithm using wrapper model and fisher score is introduced. The main objective of this paper is to prove that new hybrid model produces better accuracy than the traditional random forest algorithm.

PROBLEM STATEMENT

Predicting credit worthiness of an individual for optimized lending at the right amount of loan disbursed, interest charged and repayment window.

ADVANTAGES

1. The accuracy level of this new algorithm in finding the potential of the customer is much higher than the data mining classification algorithm and thus it proves to be very helpful for bank officers.
2. The new hybrid model produces better accuracy than the traditional random forest algorithm.

DISADVANTAGES

1. The financial institutions profitability definitely depends on the accuracy of the model.

ANALYSIS OF LOAN AVAILABILITY USING MACHINE LEARNING TECHNIQUES

AUTHOR

Sharayu Dosalwar, Ketki Kinkar, Rahul Sannat, Dr Nitin Pise

METHODOLOGY

The logistic regression model is an important predictive analytics tool for detecting loan defaulters. In order to assess and forecast, data from Kaggle is acquired. Logistic Regression models were used to calculate the various performance indicators. The models are compared using performance metrics like sensitivity and specificity. In addition to checking account details (which indicate a customer's wealth), the model is significantly better because it includes variables (customer personal attributes such as age, objective, credit score, credit amount, credit period, and so on) that should be considered when correctly calculating the probability of loan default. As a result, using a logistic regression approach, the appropriate clients to target for loan issuance can be easily identified by evaluating their plausibility of loan default. The model implies that a bank should assess a creditor's other attributes, which play a critical role in credit decisions and forecasting loan defaulters, in addition to giving loans to wealthy borrowers.

PROBLEM STATEMENT

Loans, or whether customers repay or default on their loans, affect a bank's profit or loss.

ADVANTAGES

1. The bank can minimize its Non-Performing Assets by forecasting loan defaulters.

DISADVANTAGES

1. Gender and martial status, for example, do not appear to be considered.

AN APPROACH FOR PREDICTION OF LOAN APPROVAL USING MACHINE LEARNING ALGORITHM

AUTHOR

Kathe Rutika Pramod, Panhale Sakshi Dattatray, Avhad Pooja Prakash, Dapse Punam Laxman, Ghorpade Dinesh B

METHODOLOGY

A very important approach in predictive analytics is used to study the problem of predicting loan defaulters: The Logistic regression model. The data is collected from the Kaggle for studying and prediction. Logistic Regression models have been performed and the different measures of performances are computed. The models are compared on the basis of the performance measures such as sensitivity and specificity.

PROBLEM STATEMENT

A bank's profit or a loss depends to a large extent on loans i.e., whether the customers are paying back the loan or defaulting. By predicting the loan defaulters, the bank can reduce its Non-Performing Assets.

ADVANTAGES

1. The model will focus on minimizing the overall loss in investment of bad loans in order to lessen the burden passed onto individual investors.
2. The paper will also explore privacy-preserving mechanism on sensitive information provided from the borrower credit report.

DISADVANTAGES

1. In real time customers data sets may have many missing and imputed data which needs to be replaced with valid data generated by making use of the available completed data.

LOAN APPROVAL PREDICTION

AUTHOR

Shubham Nalawade, Suraj Andhe, Siddhesh Parab, Amruta Sankhe

METHODOLOGY

The system predicts on the basis of the model that has been trained using machine learning algorithms. We have even compared the accuracy of different machine learning algorithms. We got a percentage of accuracy ranging from 75-85% but the best accuracy we got was from Logistic Regression i.e., 88.70% The system includes a user interface web application where the user can enter the details required for the model to predict.

PROBLEM STATEMENT

The banks or the financial companies take a very long time for the verification and validation process and even after going through such a regress process there is no surety that whether the applicant chosen is deserving or not. To solve this problem, we have developed a system in which we can predict whether the applicant chosen will be a deserving applicant for approving the loan or not.

ADVANTAGES

1. The web application also includes a bar plot graph of the comparison of algorithms, insights of the dataset that we have used for training the model.

DISADVANTAGES

1. The drawback of this model is that it takes into consideration many attributes but in real life sometimes the loan application can also be approved on a single strong attribute, which will not be possible using this system.