In Banking Prediction of Loan Approbation Using Machine Learning

[1]Dhanashri A. Wani, [2]Adarsha S. Tajane, [3]Tejal P. Doshi, [4]Sachin B. Bhosale

[1]Student at JCOE, Kuran
[2]Student at JCOE, Kuran
[3]Student at JCOE, Kuran
[4]Assistant Professor at JCOE, Kuran

ABSTRACT:

Distribution of loan is core business part of every bank. The acceptance and rejection of loan application is direct impact on bank's revenue and profitability. The number of applications for bank loan is increasing everyday but it is tough and risky to test out manually every application. Based on some attributes, the bank needs to decide which one is best for loan approval. This project has primary aim to predict whether the particular individual is reliable or not for loan approval. In this project, the algorithmic models are performed on dataset and therefore the different measures of performances are computed. The models are compared on the premise of the performance measures like sensitivity and specificity. the ultimate results have shown that the model produce different results. Model is marginally better because it includes variables(personal attributes of consumers like age, purpose, credit history, credit amount, credit duration, income, etc.) apart from checking amount information(which shows wealth of a customer) that ought to be taken under consideration to calculate the probability of default loan correctly. Therefore, by employing a suitable approach, the proper customers to be targeted for granting loan is easily detected by evaluating their likelihood of neglect loan.

*Keywords:-*Loan prediction, Machine Learning, Data Mining, Training and Testing

INTRODUCTION

Nowadays, the loan companies permit a loan after a very long process of verification and validation. However, they still don't have assurance that the applicant can return the loan amount without facing any difficulty. In banking industry, they have many products to sell but main source of income of any bank is on its credit line. so that they can earn from interest of these loans which they credits. A bank's profit or a loss depends to an outsized extent on loans i.e. whether the loan customers are paying back the loan or defaulting. By introducing the loan defaulters, the bank can reduce its Non-performing Assets. This makes the study of this phenomenon vital. Previous research during this era has shown that there are numerous methods to check the matter of controlling loan default. But as the right predictions are vital for the maximization of profits, it's essential to study the character of the various methods and their comparison.

GOALS AND OBJECTIVES

• To train the best machine learning model to maximize the predictive capability of deeply understanding customer's profile minimizing the risk of future loan defaulters.
• To check whether the loan approval to a specific individual is safe or not.

LITERATURE REVIEW

In this paper, we tried to minimize the risk factor behind approving a loan to the safe person to save lots of bank efforts and money. The approval of a loan mostly depends on the credit history of the candidate. Those applicants who were applying for low amounts but having high income get easily approved. In this given report we have implemented Logistic regression models as Predictive analysis tools. It is used for the given problem of prediction of loan approval.
[1].H. Ramachandra, G. Balaraju, R. Divyashree and H. Patil, "Design and Simulation of Loan Approval Prediction Model using AWS Platform," 2021 International Conference on Emerging Smart Computing and Informatics (ESCI), 2021, pp. 53-56. The main goal of this paper is to implement cloud-based machine learning algorithms. This paper is divided into four sections (i) Data Prepossessing (ii) Data Filtering (iii) Decision Tree Algorithm (iv) Logistic Regression.
[2]."Prediction for Loan Approval using Machine Learning Algorithm" Author-Ashwini S. Kadam, Shraddha R. Nikam, Ankita A. Aher, Gayatri V. Shelke, Amar S. Chandgude. Year-2021 . Aim of this paper is to predict loan safety. The SVM and Naïve bayes algorithms are used. The data cleaning is performed to avoid the missing values in the data set. The paper has following sections(i) Collection of Data, (ii) Data Cleaning and (iii) Performance Evaluation.
[3]A. Gupta, V. Pant, S. Kumar and P. K. Bansal, "Bank Loan Prediction System using Machine Learning," 2020 9th International Conference System Modelling and Advancement in Research Trends (SMART), 2020. This paper allows jumping on particular applications that deserve to be approved on

a priority basis.

[4].”Prediction of Loan Defaulter-A data Science Perspective” Author- P. Maheshwari, CH .V. Narayana, year-2020.The objective of these paper is to use data science and data analytical methods to achieve a business objective. This process involves several steps like data cleaning, preparation, modelling and model evaluation.
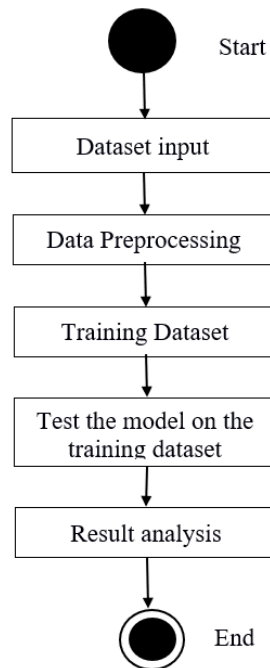
PROPOSED METHODOLOGY

Fig.(1) State Diagram

Step i: Start
Step ii: Recording the loan data(ID, married, loan amount, education, dependent Etc.)
Step iii: Data is trained.
Step iv: All Algorithms were processed on data.
**Step v:** Data is tested.

Algorithms that are used in this model are:
1. Logistic Regression
2. Decision Tree
3. Random forest

1) **Logistic Regression :** Logistic regression is supervised machine learning algorithm generally used in classification problems. In logistic regression, there are dependent variable and independent variables. Independent variables are input variables which is continuous or categorial in nature. The dependent variable are binary in nature i.e. having two categories.

2) **Decision Tree :** Decision tree also known as prediction tree which is refers to the tree structure to mention the sequences of decisions as well as consequences. It is easily possible to predict the outcome by the process of building a decision tree which has various nodes(test points) and branches. All test points demonstrate the testing specific input variable or attributes and develop a decision tree which is represented by branches. Decision trees are most probably deployed in data mining applications to solve the classification problems.

3) **Random Forest :** Random Forest is a very useful machine learning algorithm. Random Forest is a approach which uses test data for the model for training. It creates random forests for the problem set using a decision tree and then finds the solution using these random forests. At the time of the training the data Random forest creates the decision trees to predict the outcome.

SCREENSHOTS
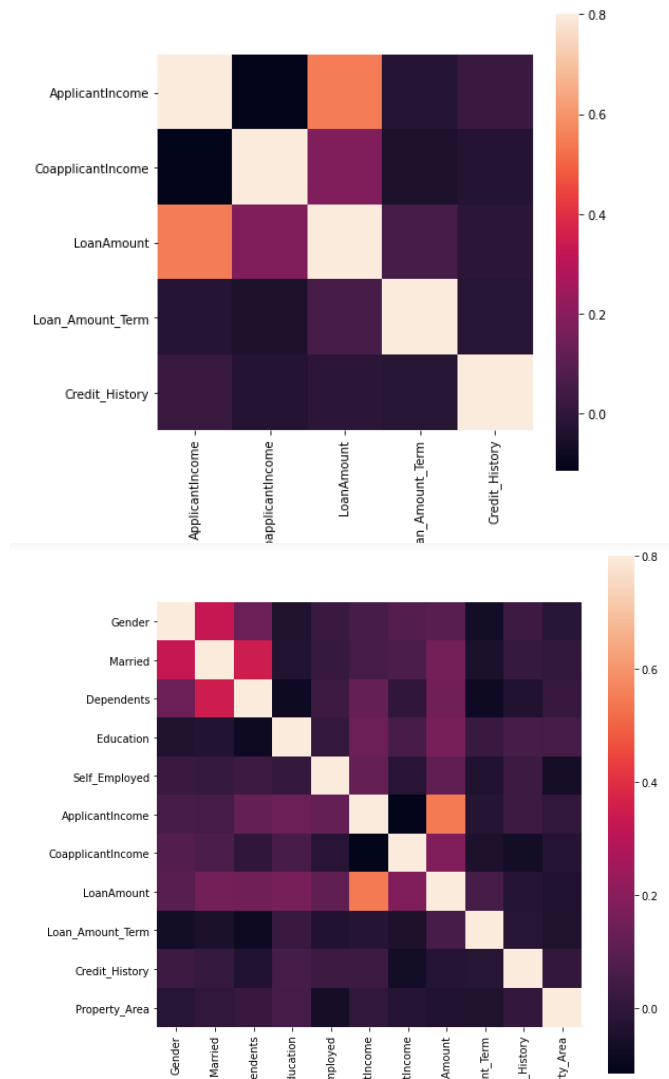
Website for collecting the information from user:

RESULT AND ANALYSIS :

There is a three machine learning approaches we have used to test the data to predict the loan defaulters from loan applications. We employed our 75% data from our available dataset for training and remaining 25% data from our dataset is employed for testing. Then prediction accuracy of the various Machine Learning approaches is calculated and compared. Based on the train data set system analyse the rest of the 25% data and predict the results in terms of loan status either accepted or rejected.

Sr. No.	Machine Learning Algorithm	Accuracy Percentage
1	Logistic Regression	83.04
2	Decision Tree	85
3	Random Forest	88.53

Table (1) Prediction accuracy of various machine learning algorithms

CONCLUSION :

From the overall discussion, it can be concluded that, the system helps us to reduce human intervention and increase efficiency of work. From the analysis of data, it is clear that it reduces the frauds done at the time of loan approval. And thus it can be said that our proposed system is efficient enough to achieved the aim of predicting the risk of future loan defaulters and can be brought into real time. Data mining strategies are helpful to the financial part for better focusing on and procuring new clients, most significant client maintenance, programmed credit endorsement which is utilized for extorsion avoidance, misrepresentation identification progressively, giving section based item, investigation of the client, exchange designs after some time for better maintenance and relationships, hazard the executives and showcasing.

REFERENCES :

[1] H. Ramachandra, G. Balaraju, R. Divyashree and H. Patil, "Design and Simulation of Loan Approval Prediction Model using AWS Platform," 2021 International Conference on Emerging Smart Computing and Informatics (ESCI), 2021, pp. 53-56, doi: 10.1109/ESCI50559.2021.9397049.

[2] A. Gupta, V. Pant, S. Kumar and P. K. Bansal, "Bank Loan Prediction System using Machine Learning," 2020 9th International Conference System Modeling and Advancement in Research Trends (SMART), 2020, pp. 423- 426, doi: 10.1109/SMART50582.2020.9336801.

[3] M. Alaradi and S. Hilal, "Tree-Based Methods for Loan Approval," 2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI), 2020, pp. 1-6, doi: 10.1109/ICDABI51230.2020.9325614.

[4] M. A. Sheikh, A. K. Goel and T. Kumar, "An Approach for Prediction of Loan Approval using Machine Learning Algorithm," 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), 2020, pp. 490-494, doi: 10.1109/ICESC48915.2020.9155614.

[5] P. Maheswari and C. V. Narayana, "Predictions of Loan Defaulter - A Data Science Perspective," 2020 5th International Conference on Computing, Communication and Security (ICCCS), 2020, pp. 1-4, doi: 10.1109/ICCCS49678.2020.9277458.