



**IJIRCCCE**

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 10, Issue 12, December 2022

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA

**Impact Factor: 8.165**



9940 572 462



6381 907 438



ijircce@gmail.com



www.ijircce.com

# Loan Approval System Using Machine Learning

K.Tejanya<sup>1</sup>, B.Preethi Mahadevi<sup>2</sup>, Amrutha S.Pattar<sup>3</sup>, Arunkumar<sup>4</sup>, Prof.Smt Gnaneshwari G.R

Dept. of CSE, Government Engineering College, Gangavathi, India

**ABSTRACT:** In India, Everyday a large number of people make application for loans for variety of purposes. The bank employees are not able to analyse or predict whether the customer can playback the amount or not for the interest rate. The aim of this Paper is to provide quick, immediate and easy way to choose the deserving applicants. It can provide special advantages to the bank. This model can be used by the bank's in making the right decision to approve or reject the loan request of the customers. A time limit can be set for the applicant to check whether his/her loan can be sanctioned or not.

## I. INTRODUCTION

The term banking can be defined as receiving and protecting money that is deposited by the individual or the entities. Everyday lots of people applying for loan in the banking sector but Bank would have limited funds. In this case, the right prediction would help the banks in predicting the funds that can be maintained. By using some class-function algorithm like logistic regression, random forest classifier, support vector machine classifier, etc for the right prediction. A Bank's profit or loss depends on the amount of loans that is whether the client is paying back the loan to the bank in right time. Recovery of loan is very important for the banking sector. In this we use machine to support our lives and make easy for us to get the candidate/client previous proofs/backup before approval of the loan amount. The Application of the client is approved or not depends on the historical data of the client by the system. Banking sector is regulated in most of the countries as it is the important factor in determining the financial stability of the country. The primary objective of the bank is to provide wealth in safer hands.

## II. LITERATURE SURVEY

In Paper [1], "Prediction for Loan Approval using Machine Learning Algorithm". In our banking system, banks have many products to sell but main source of income of any banks is on its credit line. So they can earn from interest of those loans which they credits. 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. This makes the study of this phenomenon very important. Previous research in this era has shown that there are so many methods to study the problem of controlling loan default. But as the right predictions are very important for the maximization of profits, it is essential to study the nature of the different methods and their comparison. A very important approach in predictive analytics is used to study the problem of predicting loan defaulters (i) Collection of Data, (ii) Data Cleaning and (iii) Performance Evaluation. Experimental tests found that the Naïve Bayes model has better performance than other models in terms of loan forecasting.

In Paper [2], "An Approach for Prediction of Loan Approval using Machine Learning Algorithm". In our banking system, banks have many products to sell but main source of income of any banks is on its credit line. So they can earn from interest of those loans which they credits. 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. This makes the study of this phenomenon very important. Previous research in this era has shown that there are so many methods to study the problem of controlling loan default. But as the right predictions are very important for the maximization of profits, it is essential to study the nature of the different methods and their comparison. 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. The final results have shown that the model produce different results. Model is marginally better because it includes variables (personal attributes of customer like age, purpose, credit history, credit amount, credit duration, etc.) other than checking account information (which shows wealth of a customer) that should be taken into account to calculate the probability of default on loan correctly. Therefore, by using a logistic regression approach, the right customers to be targeted for granting loan can be easily detected by evaluating their likelihood of default on loan. The model concludes that a bank should not only target the rich customers for granting loan but it should assess the other attributes of a customer as well which play a very important part in credit granting decisions and predicting the loan defaulters.

In Paper [3], “An exploratory Data Analysis for Loan Prediction based on nature of clients”. In India, the number of people applying for the loans gets increased for various reasons in recent years. The bank employees are not able to analyse or predict whether the customer can payback the amount or not (good customer or bad customer) for the given interest rate. The aim of this paper is to find the nature of the client applying for the personal loan. An exploratory data analysis technique is used to deal with this problem. The result of the analysis shows that short term loans are preferred by majority of the clients and the clients majorly apply loans for debt consolidation. The results are shown in graphs that helps the bankers to understand the client’s behavior.

In Paper[4], “Accurate Loan Approval Prediction Based on Machine Learning Approach”. Loan approval is a completely critical manner for banking groups. Banking industry continually wishes a greater correct predictive modeling device for plenty problems. Predicting credit defaulters is a tough mission for the banking industry. The system authorized or rejects the loan programs. healing of loans is a primary contributing parameter in the monetary statements of a financial institution. it's far very hard to are expecting the possibility of mortgage through the consumer. gadget studying (ML) techniques are very beneficial in predicting effects for big amount of facts. in this paper three machine studying algorithms, Logistic Regression (LR), selection Tree (DT) and Random forest (RF) are implemented to predict the loan approval of customers. The experimental results finish that the accuracy of selection Tree machine gaining knowledge of algorithm is better in comparison to Logistic Regression and Random woodland system mastering approaches.

In Paper [5], “Prediction of Modernized Loan Approval System Based on Machine Learning Approach”. This paper applied machine learning in prediction of loan approval. support Vector device (SVM) is used to expect the loan approval status of clients for financial institution loans. The consequences proven that the prediction accuracy is train Accuracy: 80% and test Accuracy: 82%. consistent with this research paper prediction accuracy is ideal when compared to different current structures. In some situations like client going through a few disaster so right here the algorithm can't expect the right end result. This research paper can find out the client is capacity or now not. In future a rigorous evaluation of different machine learning algorithms aside from these can also be achieved in future to look at the power of machine learning algorithms for loan approval prediction. It uses Artificial neural networks, Supervised learning, Unsupervised learning, Data mining.

|   |                |   |   |   |
|---|----------------|---|---|---|
| 1 | Ashwinis.kadam | Prediction for loan approval using machine learning | Existing system are often relatively inaccurate | The SVM and naïve bayes algorithm are used to predict the loan safety. Predictive analytics is used to study the problem of predicting loan defaulters. |
|---|----------------|---|---|---|

|   |                       |  |  |  |
|---|-----------------------|--|--|--|
| 2 | Mohammad ahmad sheikh | An approach for prediction of loan approval using machine learning algorithm | Calculations can get very complex,if many values are uncertain | Predictive analytic is used to study the problem of predicting loan defaulters.logistic regression model is also used for easily detection of granting loan by evaluating their likelihood of default on loan. |
|---|-----------------------|--|--|--|

|   |                |  |  |  |
|---|----------------|--|--|--|
| 3 | X.frencisjensy | An exploratory data analysis for loan prediction based on nature of clients      | Existing system often involves higher time to train the model and generally leads to the overfitting of data   | Eda is used for process of normalisation,missing value treatment,choosing essential column and visualising the data in graphical format.python is used for easy and efficient processing of data.pandas library is also used.matplot library is also used. |
| 4 | J.tejaswini    | Accurate loan approval prediction based on machine learning approach             | It is relatively expensive as the complexity and time has taken are more and little bit of noise can make it unstable which leads to wrong predictions | Six machine learning classification models have been use for prediction of andriod applications. The models are decision tree, random forest, support vector machine(SVM), linear model(LM),neural network,adaboost,pandas                                 |
| 5 | Vasantha S     | Prediction of modernized loan approval system based on machine learning approach | It takes more time to give the output  | It uses Artificial neural networks, Supervised learning, Unsupervised learning, Data mining  |

### III. PROPOSED SYSTEM

- L This proposed model will characterize the behavior of customers on the basis of their record.these records are taken from the customers,and create a data set. with the help of these data sets and training machine learning model,we predict that the customer’s loan will pass or not.

- L The aim of this paper is to provide quick,immediate and easy way to choose the deserving applicants.it can provide special advantages to the bank.the loan prediction system allows jumping to specific application so that it can be check on priority basis.

#### IV. ADVANTAGES OF PROPOSED SYSTEM

- The advantage of this system is that we provided some conditions by setting the algorithms and just by evaluating the details, we get to know eligibility criteria that client is eligible or not.
- The proposed system also scales relatively well to high dimensional data.
- The proposed system is relatively memory efficient.
- The risk of over-fitting is less in our proposed system.

#### V. METHODOLOGY

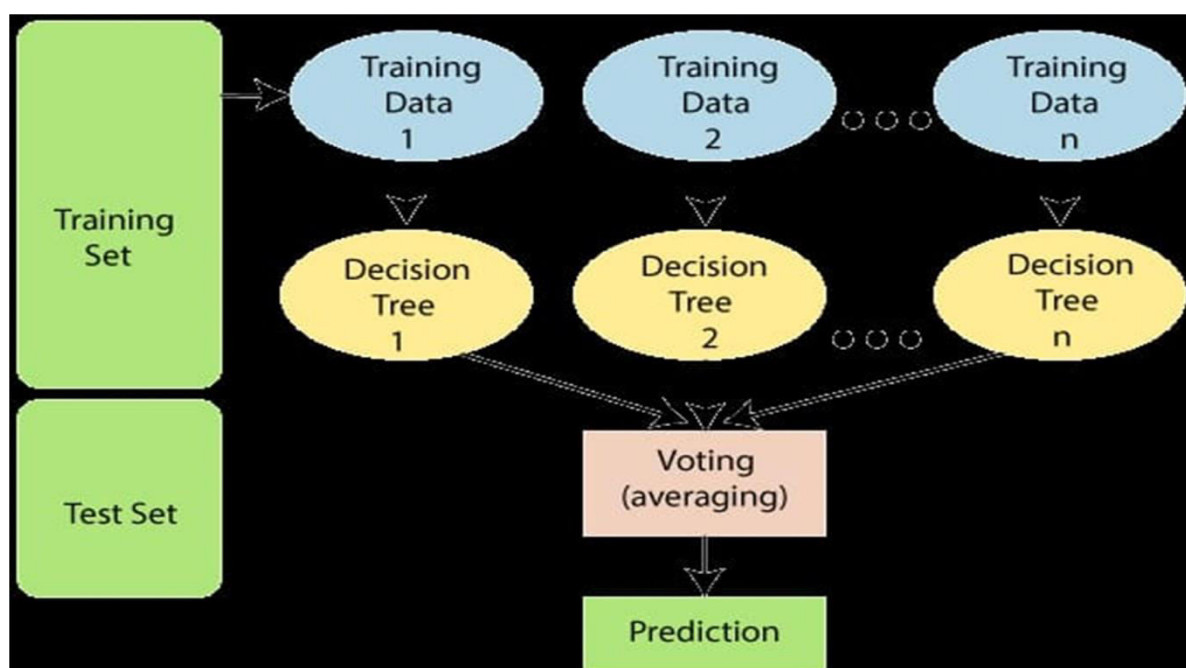
##### Algorithm:

XGBoost – XgBoost stands for Extreme Gradient Boosting, which was proposed by the researchers at the University of Washington.

- It is a library written in C++ which optimizes the training for Gradient Boosting.
- XGBoost is a Decision tree based open source software library. It implements machine learning algorithms that uses a gradient boosting framework.
- It works on Linux, Windows, and macOS.

Random Forest– Random forests is a classification algorithm which builds big number of Decision tree, whose prediction is more accurate than any of individual decision tree.

- Random Forest is a popular machine learning algorithm that belongs to the supervised learning technique.
- It can be used for both Classification and Regression problems in ML.
- As the name suggests, "Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset."
- Instead of relying on one decision tree, the random forest takes the prediction from each tree and based on the majority votes of predictions, and it predicts the final output.
- The greater number of trees in the forest leads to higher accuracy and prevents the problem of overfitting.
- The Below Diagram explains the working of the Random Forest algorithm:



**Decision Tree** –A Decision tree is a flowchart-like tree structure, where each internal node denotes a test on an attribute, each branch represents an outcome of the test, and each leaf node (terminal node) holds a class label.

- A tree can be “learned” by splitting the source set into subsets based on an attribute value test. This process is repeated on each derived subset in a recursive manner called recursive partitioning.
- The recursion is completed when the subset at a node all has the same value of the target variable, or when splitting no longer adds value to the predictions.
- A Decision tree split the dataset in to smaller parts. And then predict the every chances.

**SYSTEM REQUIREMENTS :**

**Data Collection:** This is the first real step towards the real development of a machine learning model, collecting data. This is a critical step that will cascade in how good the model will be, the more and better data that we get, the better our model will perform.

**Dataset:** The dataset consists of 511 individual Married : Applicant married (Y/N) Application\_Status: (Target) data. There are 10 columns in the dataset which are described below.

Application\_ID : Unique Loan ID

Gender : Male/ Female

Self\_Employed :Self employed (Y/N)

Married : Applicant married (Y/N)

Dependents : Number of dependents

Education : Applicant Education (Graduate/ Under Graduate)

Credit\_History : credit history meets guidelines



Property\_Area : Urban/ Semi Urban/ Rural

Income : Applicant income

Application\_Status: (Target) Loan approved (Y/N)

Data Preparation: Wrangle data and prepare it for training. Clean that which may require it (remove duplicates, correct errors, deal with missing values, normalization, data type conversions, etc.).

## VI. FUTURE ENHANCEMENTS

In future a rigorous analysis of other machine learning algorithms other than these can also be done in future to investigate the power of machine learning algorithms for loan approval prediction.

## VII. CONCLUSION

The project applied machine learning in prediction of loan approval. Support Vector Machine (SVM) is used to predict the loan approval status of customers for bank loans. The results shown that the prediction accuracy is Train Accuracy: 80% and Test Accuracy: 82%. According to this research paper prediction accuracy is good when compared to other existing systems. In some situations, like client going through some disaster so here the algorithm cannot predict the appropriate result. This research paper can find out the client is potential or not.

## REFERENCES

- [01].Ashwini S. Kadam, Shraddha R Nikam, Ankita A. Aher, Gayatri V. Shelke, Amar S. Chandgude. "Prediction for Loan Approval using Machine Learning Algorithm", Apr 2021 International Research Journal of Engineering and Technology (IRJET).
- [02]. Mohammad Ahmad Sheikh, Amit Kumar Goel, Tapas Kumar. "An Approach for Prediction of Loan Approval using Machine Learning Algorithm", 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), 2020.
- [03].X.FrencisJensy, V.P.Sumathi,Janani Shiva Shri, "An exploratory Data Analysis for Loan Prediction based on nature of clients", International Journal of Recent Technology and Engineering (IJRTE), Volume-7 Issue-4S, November 2018.
- [04]. J. Tejaswini<sup>1</sup>, T. MohanaKavya, R. Devi Naga Ramya, P. SaiTriveniVenkataRaoMaddumala. "ACCURATE LOAN APPROVAL PREDICTION BASED ON MACHINE LEARNING APPROACH" Vol 11, [www.jespublication.com](http://www.jespublication.com), page 523 Issue 4, April/ 2020 ISSN NO: 0377-9254.
- [05]. Vasantha S, SoumyaRanjan Jena. "PREDICTION OF MODERNIZED LOAN APPROVAL SYSTEM BASED ON MACHINE LEARNING APPROACH" VOL 04, [www.irjmets.com](http://www.irjmets.com), Issue:06/June-2022.



**INNO**  **SPACE**  
SJIF Scientific Journal Impact Factor

**Impact Factor: 8.165**

**doi**<sup>®</sup>  
**cross** **ref**

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
**INDIA**



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

 **9940 572 462**  **6381 907 438**  **ijircce@gmail.com**



[www.ijircce.com](http://www.ijircce.com)

Scan to save the contact details