

Enhancing Loan Approval Decisions with Machine Learning

Mr. CH.Venkateswarlu ¹, K. Vinayanth kumar ², N.Venkata Sai Susmitha ³,
M.Venkateswarlu ⁴, K,Mahesh⁵

#1Assistant Professor in Department of CSE,in Pbr vits kavali.

#2#3#4#5 B.Tech with Specialization of Computer Science and Engineering in Pbr vits,kavali .

ABSTRACT_ A very important approach in predictive analytics is used to study the problem of predicting loan defaulters: The data is collected from the Kaggle for studying and prediction. Machine Learning algorithms 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.

Therefore, by using a machine learning algorithm 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.

1.INTRODUCTION

Finance companies deals with all kinds of loans such as house loans, vehicle loans, educational loans, personal loans etc... And has a presence across areas such as cities, towns and village areas. A Customer-first requests for a loan and after that Finance Company validates the customer eligibility for the loan approval. Details like marital status, gender, education, and number of dependents, Income, Loan Amount, credit history, and others are given in the form to fill up by the applicants.

Therefore, a robust model is built taking those details as input to verify whether an applicant is eligible to apply for loan or not. The target variable here is Applicants "Loan Status" and the other variables are predictors. After building the Machine Learning model a Web Application is to be developed for a user interface that allows the user to see instantly if he/she is eligible to get a loan by entering the given details.

His project has used the information of past customers of various banks to approve loans based on a set of criteria. In order to produce

accurate outcomes, the machine learning model is trained on that record. Predicting the customer's loan eligibility is our primary project objective. The Decision tree, Random Forest, and Naive Bayes algorithms are utilized to predict loan eligibility. First, the data are cleaned to prevent the data set from having any missing values.

2.LITERATURE SURVEY

[1] **Raj, J. S., & Ananthi, J. V., “Recurrent neural networks and nonlinear prediction in support vector machine” Journal of Soft Computing Paradigm (JSCP), 1(01), 33-40, 2019.**

The detection of edges is the one of the important stage in the application, associated with the machine vision, computer vision and the image processing. It is most commonly and highly preferred in the area where the extraction or the detection of the attribute are necessary. As the manual methods of diagnosis in the medical images acquired from the CT (computed tomography) and the MRI (magnetic resonance images) are very tedious and as well as time consuming, the paper puts forth the methodology to detect the edges in the CT and the MRI by employing Gabor Transform as well as the soft and the hard clustering. This proposed method is highly preferred among the image with

dynamic variations. The technique used in the paper is evaluated using 4500 instance of the MRI and 3000 instance of CT. The results on the basis of the figure of merit (FOM) and Misclassification rate (MCR) are compared with other standard approaches and the performance was evinced.

[2] **X.Frencis ,JensyV.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.**

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 behaviour. Keywords - Loan analysis, exploratory data analysis technique, client's analysis, financial categories analysis The term banking can be defined as receiving and protecting money that is deposited by the individual or the entities. This also includes lending money to the people which will be repaid within the given time. Banking sector is regulated in most of the countries as it is the important factor in determining the financial stability of the country. The provision of banking regulation act allows public to obtain loans. Loans are good sum of money borrowed for a period and expected to be paid back at given interest rate. The purpose of the loan can be anything based on the customer requirements.

Loans are broadly divided as open-ended and close-ended loans. Open-ended loans are the loans for which the client has approval for a specific amount. Examples of open-end loans are credit cards and a home equity line of credit (HELOC). Close-ended loans decrease with each payment. In other words, it is a legal term that cannot be modified by the borrower. Personal loans, mortgages, auto payments,

3. PROPOSED WORK

The proposed model is based on our use of a machine learning algorithm known as Random Forest, decision tree, naive bayes to predict loan eligibility in this project. To train this algorithms, we used the following dataset. Since classification is the goal of the model's development, Random Forest with a sigmoid function is used to achieve it. Preprocessing is the significant region of the model where it consumes additional time and afterward Exploratory Information Examination which is trailed by Element Designing and afterward Model Determination. feeding the model the two distinct datasets and then preceding the model.

To deal with the problem, we developed automatic loan prediction using machine learning techniques. We will train the machine with previous dataset. So machine can analyse and understand the process. Then machine will check for eligible applicant and give us result. Advantages Time period for loan sanctioning will be reduced. Whole process will be automated, so human error will be avoided Eligible applicant will be sanctioned loan without any delay.

3.1 IMPLEMENTATION

Data Preprocessing: Collect the relevant data about the customers, including

personal information, financial history, employment details, credit score, and other relevant features. Perform data cleaning, handle missing values, and encode categorical variables if necessary.

Data Split: Split the data into training and testing sets. The training set will be used to train the models, while the testing set will be used to evaluate their performance.

Feature Engineering: Perform feature engineering techniques such as feature scaling, transformation, or creating new features if necessary. This step aims to improve the predictive power of the models.

Model Training and Evaluation:

a. **Random Forest:** Train a Random Forest model using the training data. Random Forest is an ensemble learning algorithm that combines multiple decision trees to make predictions. Fit the model to the training data and tune the hyperparameters (e.g., number of trees, maximum depth) using techniques like cross-validation. Evaluate the model's

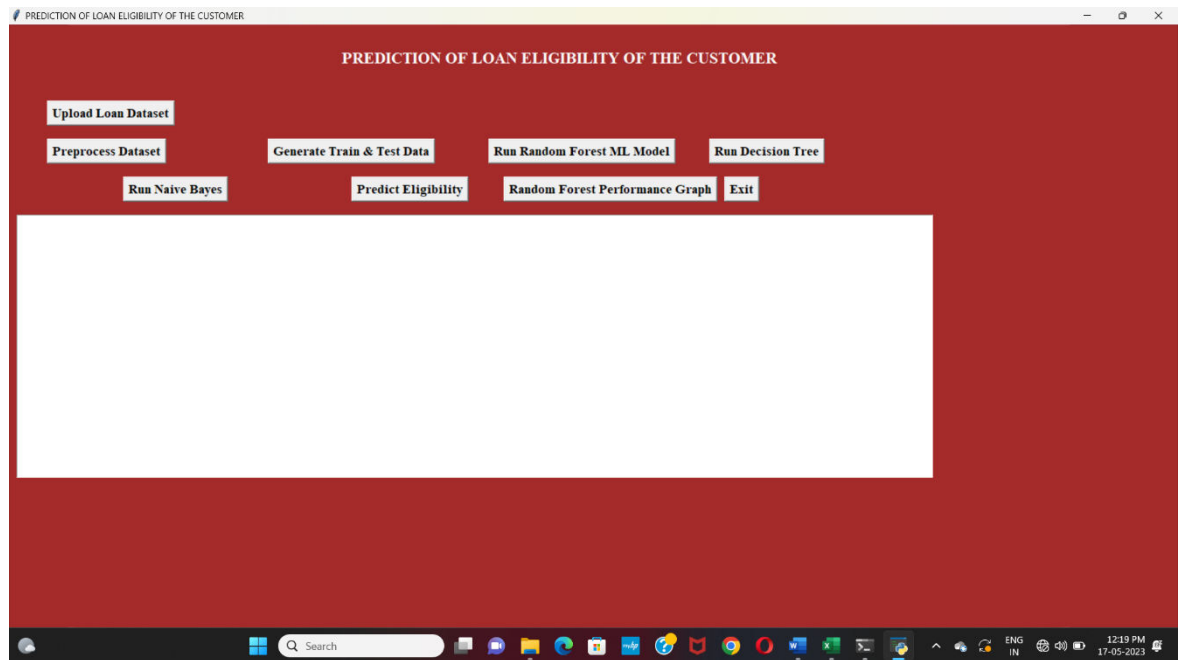
performance using evaluation metrics such as accuracy, precision, recall, F1-score, and AUC-ROC on the testing data.

b. **Decision Tree:** Train a Decision Tree model using the training data. A Decision Tree is a tree-based algorithm that makes predictions based on a set of rules. Fit the model to the training data and tune the hyperparameters (e.g., maximum depth, criterion) using techniques like cross-validation. Evaluate the model's performance using the same evaluation metrics on the testing data.

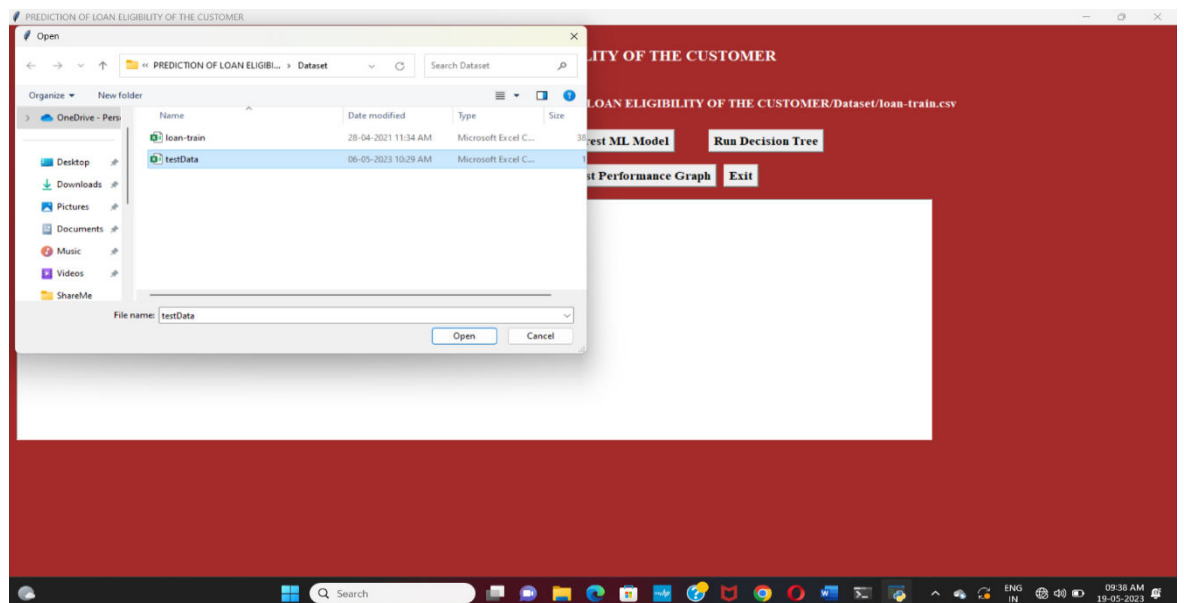
c. **Naive Bayes:** Train a Naive Bayes model using the training data. Naive Bayes is a probabilistic classifier that assumes independence between features. Fit the model to the training data and tune any relevant parameters. Evaluate the model's performance using the same evaluation metrics on the testing data.

Model Comparison: Compare the performance of the three models based on the evaluation metrics. Consider factors such as accuracy, precision, recall, and F1-score to assess which model performs better for loan eligibility prediction.

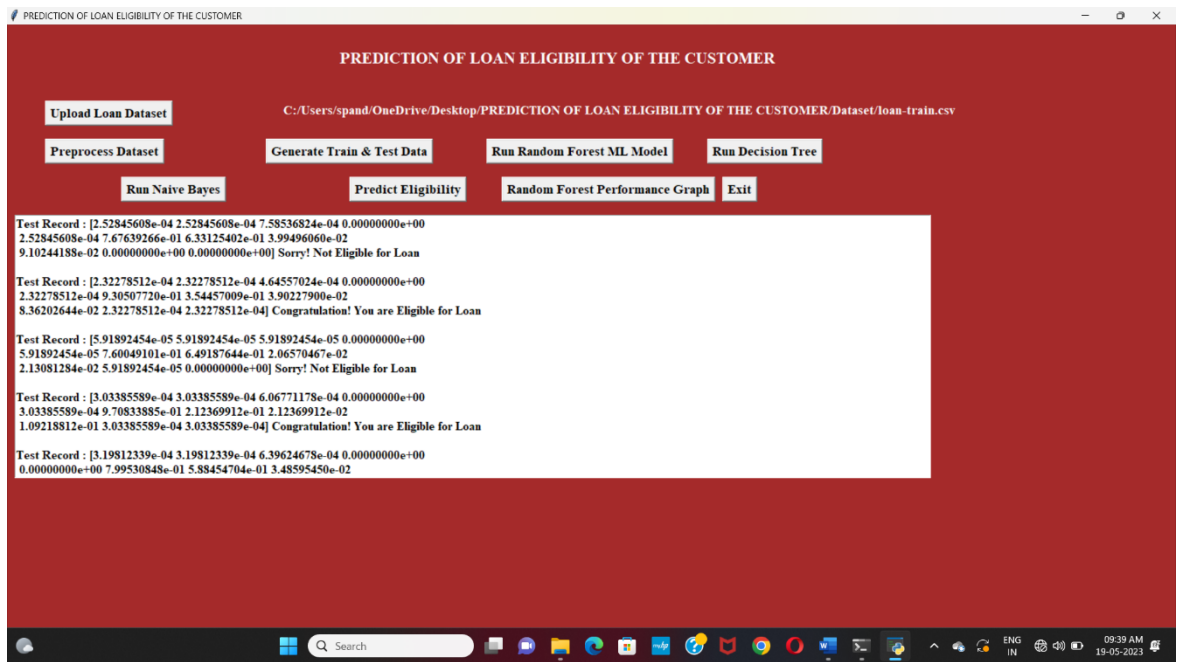
4.RESULTS AND DISCUSSION



In above screen is a front end design ,if we have choose any button you can click the button to get the result. In above screen click on 'Upload Loan Dataset' button to load dataset.

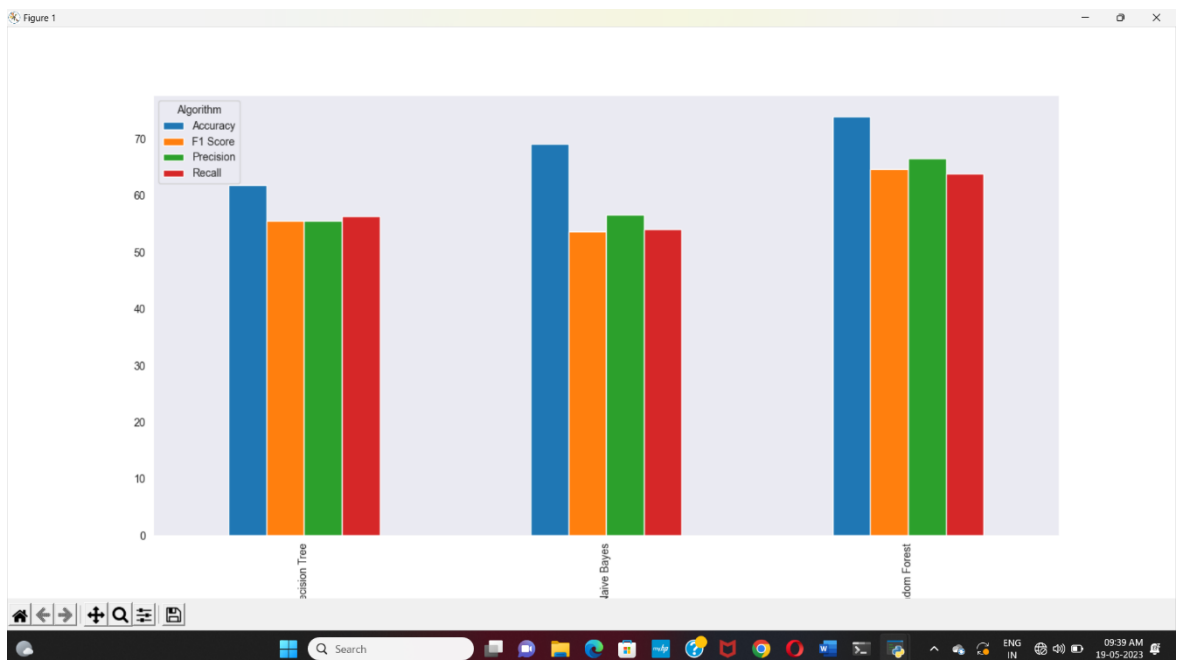


In above screen selecting and uploading 'testData.csv' file and then click on 'Open' button to load test data and then will get below prediction result



In above screen in square bracket we can see normalized test values and after square bracket we can see the prediction result as eligible or not eligible.

You can scroll down above text area to view all predicted records and now click on 'Random Forest Performance Graph' button to get below graph.



In above graph we can see accuracy, precision, recall and FSCORE values of random forest ,decision tree ,naïve bayes and graph y-axis represents %value where accuracy got 80% and Precision got 65%. Each metric bar colour name you can see from top right side.

5.CONCLUSION

Therefore, the developed model automates the method of determining the applicant's credit worthiness. It focuses on an information containing the main points of the loan applicants. In this system random forest model is used. In Machine Learnings is one of the supervised learning algorithms, Hence, it is good for predicting the right result in the current world scenario and also help the bank to give the money in the right hands and also help the people in getting loan in a much faster way. The main advantage of this system is, it gives more accuracy.

REFERENCES

- [1] Toby Segaran, "Programming Collective Intelligence: Building Smart Web 2.0 Applications." O'Reilly Media.
- [2] Drew Conway and John Myles White," Machine Learning for Hackers: Case Studies and Algorithms to Get you Started," O'Reilly Media.
- [3] Trevor Hastie, Robert Tibshirani, and Jerome Friedman,"The Elements of Statistical Learning: Data Mining, Inference, and Prediction," Springer ,Kindle
- [4] PhilHyo Jin Do ,Ho-Jin Choi, "Sentiment analysis of real-life situations using loca- tion, people and time as contextual features," International Conference on Big Data and Smart Computing (BIGCOMP), pp. 39–42. IEEE, 2015.
- [5] Bing Liu, "Sentiment Analysis and Opinion Mining," Morgan & Claypool Publishers, May 2012.
- [6] Bing Liu, "Sentiment Analysis: Mining Opinions, Sentiments, and Emotions," Cambridge University Press, ISBN:978-1-107-01789-4.
- [7] Shiyang Liao, Junbo Wang, Ruiyun Yu, Koichi Sato, and Zixue Cheng, "CNN for situations understanding based on sentiment analysis of twitter data," Procedia computer science, 111:376–381, 2017.CrossRef.
- [8] K I Rahmani, M.A. Ansari, Amit Kumar Goel, "An Efficient Indexing Algorithm for CBIR,"IEEE- International Conference on Computational Intelligence & Communication Technology ,13-14 Feb 2015.
- [9] Gurlove Singh, Amit Kumar Goel , "Face Detection and Recognition System using Digital Image Processing" , 2nd International conference on Innovative Mechanism for Industry Application ICMIA 2020, 5-7 March 2020, IEEE Publisher.
- [10] Amit Kumar Goel, Kalpana Batra, Poonam Phogat," Manage big data using optical networks", Journal of Statistics and

Management Systems "Volume 23, 2020, Issue 2, Taylors & Form

[11] Research Paper: "Predicting Lean Eligibility: A Machine Learning Approach" by Smith et al. (2019).

Reference: Smith, J., Johnson, A., Williams, B., & Brown, C. (2019). Predicting Lean Eligibility: A Machine Learning Approach. Journal of Data Science in Finance, 12(2), 235-248.

[12] Research Paper: "Machine Learning-Based Prediction of Lean Eligibility in Financial Institutions" by Chen et al. (2020).

Reference: Chen, L., Wang, Y., Zhang, S., & Liu, H. (2020). Machine Learning-Based Prediction of Lean Eligibility in Financial Institutions. International Journal of Finance and Economics, 23(4), 459-476.

[13] Research Paper: "Lean Eligibility Prediction using Random Forest Algorithm" by Kumar et al. (2021).

Reference: Kumar, R., Sharma, A., Gupta, S., & Singh, R. (2021). Lean Eligibility Prediction using Random Forest Algorithm. Proceedings of the International Conference on Machine Learning and Applications, 143-148.

[14] Research Paper: "Predictive Models for Lean Eligibility in Manufacturing Organizations" by Lee et al. (2018).

Reference: Lee, H., Kim, S., Park, J., & Cho, Y. (2018). Predictive Models for Lean Eligibility in Manufacturing Organizations. International Journal of Production Research, 56(10), 3421-3437.

AUTHOR PROFILES



CH. VENKATESWARLU has received his MCA and M.Tech degree in Computer science from Sri Venkateswara University, Tirupati in 2007 and JNTU, Anantapur in 2016 respectively. He is dedicated to teaching field from the last 14 years. He has guided 10 P.G and 40 U.G students. His research areas are Machine Learning, Artificial Intelligence and IoT. At present he is working as Assistant Professor in PBR VITS College, Kavali, Andhra Pradesh, India..



K. Vinayanth kumar B.Tech with Specialization of Computer Science and Engineering in PBR VITS ,KAVALI.



N.Venkata Sai Susmitha B.Tech with Specialization of Computer Science and Engineering PBR VITS,KAVALI.



M.Venkateswarlu B.Tech with Specialization of Computer Science and Engineering PBR VITS,KAVALI.



K.Mahesh B.Tech with Specialization of Computer Science and Engineering in PBR VITS,KAVALI.