IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.918 ⁜ Vol. 12, Issue 3, March 2023

DOI: 10.17148/IJARCCE.2023.12316

Small Business Loan Approval Prediction using Machine Learning Algorithms Techniques

Nakil Saiyukta[1], Sonali Chintewar[2], Nupur Chauhan[3], Aarati Lonari[4]

Student, Department of Computer Engineering, RGCOE, Ahmednagar, India[1-4]

Abstract: By obtaining capital and making loans, banks' core business model relies on financial intermediation (mortgage, real estate, and consumer and company's loans). The latter is the biggest contributor to credit risk and is made up of two key components: loan approval and fraud. We will concentrate on loan approval using machine learning models in this post. Although banks can sell various items in our financial system, their main source of income comes from lines of credit. Therefore, they can benefit from interest on the credited loan. A bank's profit or loss is largely determined by the loans it makes, i.e. whether its customers repay the loans. Banks can reduce their non-performing assets by predicting defaults. This highlights the importance of studying this phenomenon. Based on previous research from this time, there are several ways to study the problem of default prevention. But since accurate predictions are key to maximizing returns, understanding how the different methods of work.

Keywords: Big Data, Machine Learning, Python, Logistic Regression, SVM, Decision Tree, Naive Bayes, Loan Prediction.

I. INTRODUCTION

While approving loans, banks are having a lot of trouble. The bank workers oversee a large number of applications every day, which is difficult and increases the likelihood of mistakes. The majority of banks make money through loans,however it can be dangerous to select eligible clients from the pool of applicants. A bank could suffer a severe loss due to one error. Almost all banks' main line of operation is loan distribution. This study paper used information on prior clients of different banks whose loans were accepted according to a set of conditions. In order to obtain reliable results, based on the record, the machine learning model is trained. Predicting loan safety is the primary goal of this study. The logistic regression approach is used to forecast loan safety. The data is first cleaned to remove any missing values from the data collection. To predict if a customer's loan application will be approved, we would discuss the Decision Tree algorithm, Random Forest method, Logistic Regression algorithm, SVM algorithm, and KNN algorithm are five machine learning algorithms.

Our main objective is to apply machine learning ideas to determine a customer's loan status and forecast a prompt, exact result that helps the lender analysis the situation, improve services, and reduce risk by choosing the right candidate, saving the time, work and money. Also, we would evaluate different machine learning algorithms and choose the best one.

II. LITERATURE SURVEY

Ritika Purswani and Sakshi Verma author proposed that the system able to quickly, conveniently, and effectively select the most worthy applicants satisfies & it is very useful to the loan prediction system. That could offer the bank special advantages. We have examined how to create a Loan Approval Prediction System in this research. The analytical steps in developing this system include data gathering, exploratory data analysis, data preprocessing, model construction, and model testing [1].

Nitesh Pandey, Ramanand Gupta Machine learning was applied in this study to forecast loan approval. The first steps in the prediction approach are the pre-processing of the data, filling in the missing values, and the analysis of experimental data. For each of these methods, a precision rate of between 70% and 80% after being evaluated on the test dataset. However it is apparent that the Support Vector Machine model in this situation is more effective and yields better outcomes than other models [2].

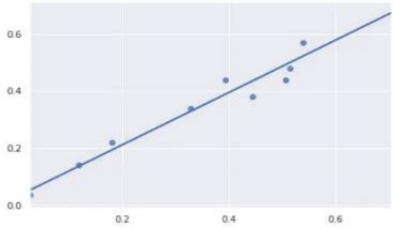
Mohammad Ahmad Sheikh, Amit Kumar Goel The steps in the prediction process include the Data cleansing and processing, imputation of missing values, experimental analysis of the data set, model creation, model evaluation, and model testing. Based on the original data set, the data set's greatest degree of accuracy is 0.811.After considering that applicants with the lowest credit scores won't be approved for loans because they have a higher likelihood of not repaying the loan amount, the following conclusions are drawn. It makes sense that applicants who request smaller loans but more

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

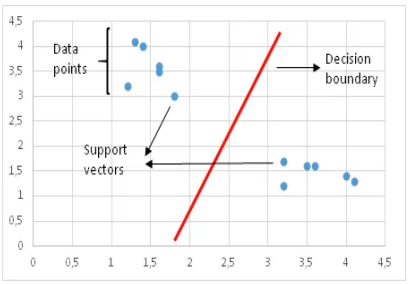
ISO 3297:2007 Certified ⁜ Impact Factor 7.918 ⁜ Vol. 12, Issue 3, March 2023

DOI: 10.17148/IJARCCE.2023.12316



likely to be accepted since they are also more likely to pay back their debts if they have high earnings. Other aspect, such as gender and marital status, don't seem to be taken into consideration [3].

Anant Shinde, Yash Patil, author suggest that prediction model uses a logistic regression technique. A logistic classification model that predicts loan status was created using approximately 600 sample data that were collected and evaluated. This precision is attained using regression models, and the procedure has a maximum accuracy of about 82 percent. The model is quick to adjust to a wide range of inputs and can predict outcomes. Also, this tactic saves the financial sector's workers a substantial amount of time [4].

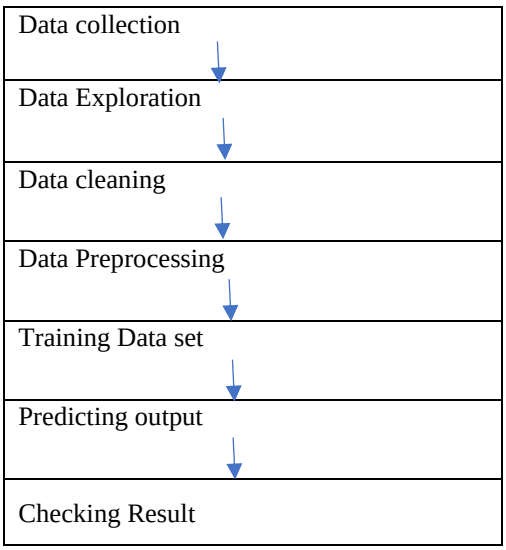


Kanishk Guptaa, Binayak Chakrabartib This methodology will help lenders and the financial sector as a whole avoid the negative effects of consumer debt turning into a non-performing asset. The development of predictive algorithms, such as machine learning, has aided in addressing some of society's most pressing problems [5].

Vipashi Kansal, Dr.Upma Jain Author used a variety of machine learning concepts, including data preparation, dataset splitting, dataset balancing, training and testing numerous classification models, and comparing the accuracy of those models under various conditions. These techniques are effective at determining if a consumer will be approved for a loan. Our study can aid in understanding various machine learning algorithms and various techniques to take [6].

Kumar Arun, Garg Ishan, Kaur Sanmeet author represented that the system is safe to say that After proper consideration of the strengths and limitations of the components, the product is a very effective component. The app works well and meets all the required specifications. The
Component easily connects to many other systems [7].

III. PREDICTION LOAN METHODOLOGY



1. Linear Regression

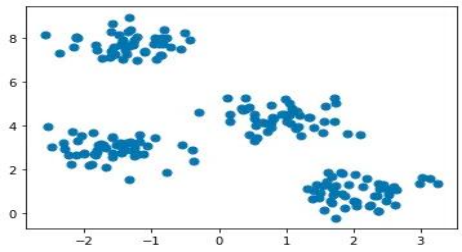
Supervised learning algorithms for linear regression attempt to describe the relationship between a continuous target variable and one or more independent variables by fitting a linear equation to the data.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.918 ⁜ Vol. 12, Issue 3, March 2023

DOI: 10.17148/IJARCCE.2023.12316

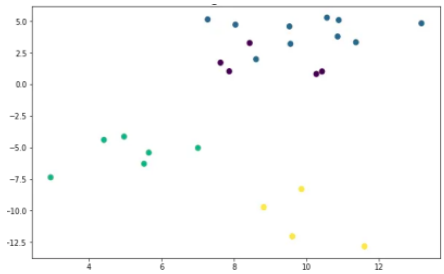


2. Support Vector Machine

The Support Vector Machine (SVM) learning algorithm is commonly used for classification tasks, but can also be used for regression tasks. SVM creates boundary decisions to Differentiate between classes. The most Important aspect of the SVM algorithm is How the decision boundary is drawn or determined. Before creating the decision boundary, each observation (or data point) is plotted in n-dimensional space. The number of features used is "n". 3.

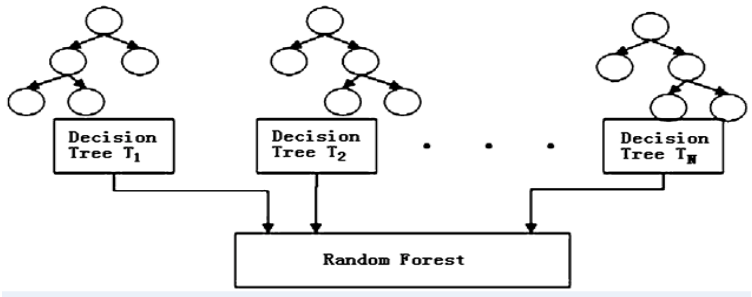
3. K-Nearest Neighbors (kNN)

A supervised learning method called K-Nearest Neighbors (kNN) can be used to solve classification and regression problems. The basic principle of kNN is that the value or class of a data point is determined by nearby data points. The class of a data point is determined by a kNN classifier based on majority voting.



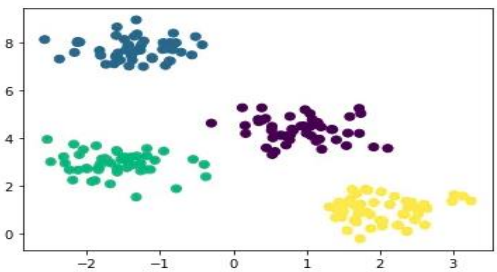
4. Random Forest (RF)

A random forest is a collection of many decision trees. The bagging technique is used to build random forests, while decision trees are utilised as parallel estimators. When applied to a classification problem, the solution is decided by the findings from the majority of decision trees. The prediction in a regression is the average value of the target values in a leaf node.The mean value of the decision tree outcomes is taken into account by random forest regression.



5. K-Means Clustering

In this approach to arrange a collection of data points so that related data points are combined together is by clustering. Because of this, clustering algorithms search for similarities or differences between data points. Since clustering is an unsupervised learning method, data points have no labels attached to them. Clustering techniques look for the data's underlying structure.



International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁜ Impact Factor 7.918 ⁜ Vol. 12, Issue 3, March 2023
DOI: 10.17148/IJARCCE.2023.12316

IV. CONCLUSION

This model will be used by anyone wishing to apply for a loan or through the banking system. The study of the data becomes clear that it reduces all frauds committed during the loan approval process. Everyone values their time, thus by doing this, the wait time for both the bank and the applicant will be reduced. Data cleaning and processing, missing value imputation, experimental analysis of the data set, model construction, and model testing on test data are all steps in the prediction process. In the future, this prediction module may be improved and incorporated. The system is trained using prior training data, but it is feasible to alter the software in the future so that it may accept new testing data as well as training data and predict as necessary.

REFERENCES

[1] Ritika Purswani, Sakshi Verma Loan Approval Prediction using Machine Learning: A Review © 2021, IRJET Volume: 08 Issue: 06 | June 2021
[2] Nitesh Pandey1, Ramanand Gupta2 "Loan Approval Prediction using Machine Learning Algorithms Approach" © June 2021| IJIRT | Volume 8 Issue 1 | ISSN: 2349-6002
[3] Mohammad Ahmad Sheikh, Amit Kumar Goel "An Approach for Prediction of Loan Approval using Machine Learning Algorithm" 978-1-7281-4108-4/20/$31.00 ©2020 IEEE
[4] Anant Shinde1, Yash Patil2, Ishan Kotian3, Abhinav Shinde4and Reshma Gulwani5 " Loan Prediction System Using Machine Learning "ITM Web of Conferences 44, 03019 (2022) https://doi.org/10.1051/itmconf/20224403019 ICACC-2022
[5] Kanishk Guptaa,*, Binayak Chakrabartib Loanification - Loan Approval Classification using Machine Learning Algorithms ICICC-2021
[6] Vipashi Kansal, Dr.Upma Jain, Ashutosh Kumar Gupta, Ms Manisha Aeri" Home Loan Prediction Using Machine Learning Models"Ilkogretim Online - Elementary Education Online, 2021; Vol 20 (Issue 3): pp. 3442-3449 http://ilkogretim-online.org doi: 10.17051/ilkonline.2021.03.359
[7] Loan Approval Prediction based on Machine Learning Approach Kumar Arun, Garg Ishan, and Kaur Sanmeet.

