

Prediction of Modernized Loan Approval System Based on Machine Learning Approach

Mr.Md.Azhar, Yenni Jyothi², Venganti Sreeja², Sasupally Krishna Kiriti², N Akhil Kumar^{2,1Asst} *Professor, Computer Science and Engineering, CMR Engineering College, medchal, T.S, India*^{2B.Tech, Computer Science and Engineering, CMR Engineering College, medchal, T.S, India}

Abstract: The quality of life for humans has improved thanks to technology. We intend to provide something fresh and unique every day. In the banking industry, the applicant receives proofs/backup prior to acceptance of the loan amount. We have machines to support our life and make us somewhat complete. The system's past data about the applicant determines whether or not the application is granted. Numerous individuals ask for loans every day in the banking industry, yet the bank's resources are constrained. A classes-function algorithm would be extremely helpful in this situation if the proper prediction could be made. Examples include support vector machine, logistic regression, and random forest classifiers. The quantity of loans a bank makes or loses depends on whether the client or customer pays back the loan. The most crucial task for the banking industry is loan recovery. The process of improvement is crucial in the banking industry. utilising several categorization techniques, a machine learning model was constructed utilising the historical data of the applicants. The major goal of this study is to use machine learning models trained on historical data to predict whether a new application will be approved for a loan or not.

1. INTRODUCTION

We can determine if a loan will be approved or not using the prediction of a modernised loan approval system based on a machine learning technique. We collect information from the user for this system, such as his monthly salary, marital status, loan amount, loan term, etc.

The bank will next determine whether the customer will get the loan based on its criteria.

Thus, a classification system exists. In this system, a training set is used to create the model, and the classifier is then able to categorise the data items into the correct classes. In order to train the data and provide the correct outcome, which is the client's potential and ability to repay the loan, a test dataset is constructed.

For banks and customers alike, the prediction of a modernised loan approval procedure is quite useful. This method evaluates each applicant based on their priority. The customer may

submit his application directly to the bank, in which case the bank will handle the whole procedure without interference from a third party or shareholder. Finally, based on its priority system, the bank will determine if the applicant is worthy or not. This study paper's only goal is to ensure that the worthy applicant receives straightforward answers right away.

A Description of The Project: Three machine learning methods are employed in this study's research to determine the data set's accurate prediction.

(a) **XGBoost** - An open source software library built on decision trees, XGBoost. It employs gradient boosting as a foundation for machine learning techniques. It runs on Windows, macOS, and Linux.

(b) **Random Forest**- Random forests is a classification technique that creates several Decision Trees, each of which has a higher predictive accuracy than the others.

2. LITERATURE SURVEY

2.1 Existing System

For financial companies, the loan approval procedure is crucial. For numerous difficulties, the banking industry is constantly in need of a more precise predictive modelling system. For the financial sector, predicting loan defaulters is a challenging issue. The loan applications were accepted or rejected by the system. A key determining factor in a bank's financial results is loan recovery. Predicting whether a consumer will pay back a debt is exceedingly tough. For vast amounts of data, machine learning (ML) methods are highly helpful in predicting outcomes.[1] Three machine learning techniques, Logistic Regression (LR), Decision Tree (DT), and Random Forest (RF), are used in the proposed system to forecast whether or not consumers would be approved for loans. [2] According to the experimental findings, the Decision Tree machine learning algorithm is more accurate than Logistic Regression and Random Forest machine learning methods.

2.2 Proposed System

This suggested model will evaluate consumer behaviour based on their past behaviour. These client records are collected to form a data set.[3] We make predictions about whether the customer's loan will be approved or not using these data sets and a machine learning model that has been trained.

This article will cover the benefit of loan prediction.[4]Machine algorithms determine whether a consumer will be able to repay the loan or not. With the use of this technique, we can determine whether or not a loan applicant will be able to repay the loan. [5] We forecast that the customer will be qualified for a loan if they can repay the amount. Additionally, we forecast that the customer is ineligible if the candidate fails.

2.2 Analysis of Loan Prediction Data

The issue of how we evaluate whether to approve the loan or not emerges. We supply the loan to our consumer based on two goal criteria. We must verify all the requirements, including evidence of income, residence, and identification. The consumer is then given the loan, whether they are qualified to repay it or not. The middle class has a significant demand for loans since parents need them for their children's education as well as for their businesses.[6]

Some individuals have abrupt financial crises, while others attempt to defraud banks of their money.

As a result, we must double-check everything since banks are not experiencing NPA loans.

Higher possibilities of loan repayment are associated with better customers.

Background checks should be thorough so that we can anticipate receiving the loan back at the ideal moment.

As a result, we analyse on a variety of bases, which are referred to as our target variables.

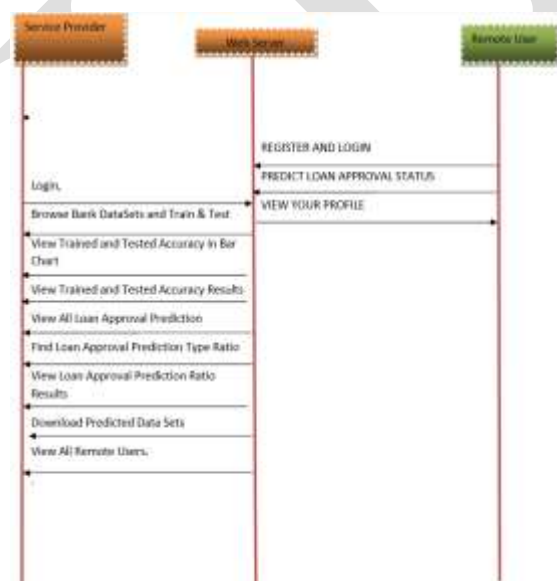


Fig 1: System Flow

3. SYSTEM ANALYSIS&DESIGN

3.1 Method of Predicting Loans

This suggested model will evaluate consumer behaviour based on their past behaviour.

These client records are collected to form a data set. We forecast whether the customer's loan will be approved or not using these data sets and a training machine learning model.

Algorithms in the machine foretell whether a consumer will be able to repay the loan or not.

3.3 System Architecture: Value of The Proposed System

We'll talk about the benefit of loan prediction in this. With the use of this technique, we can determine whether or not a loan applicant will be able to repay the loan. We forecast that the customer will be qualified for a loan if they can repay the amount.

Additionally, we forecast that the customer is ineligible if the candidate fails.

The benefit of this system is that by establishing specific conditions and defining the algorithms, we can determine whether a client meets the eligibility requirements just by looking at the facts.

This system may be created to forecast if a user's loan application will be granted by the bank based on inputs from the user such as salary, address, loan amount, and loan term.

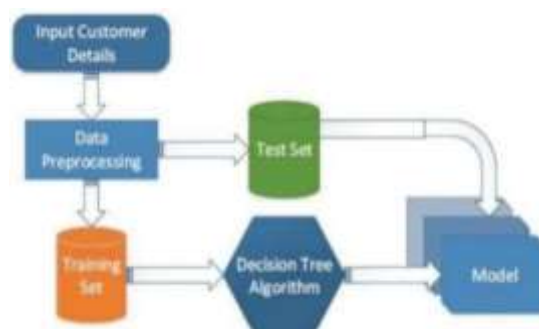


Fig 2 : System Architecture

3.4 Data Flow Diagram : Whenever a new system is developed, user training is required to educate them about the working of the system so that it can be put to efficient use by those for whom the system has been primarily designed. For this purpose the normal working of the project was demonstrated to the prospective users. Its working is easily understandable and since the expected users are people who have good knowledge of computers, the use of this system is very easy.

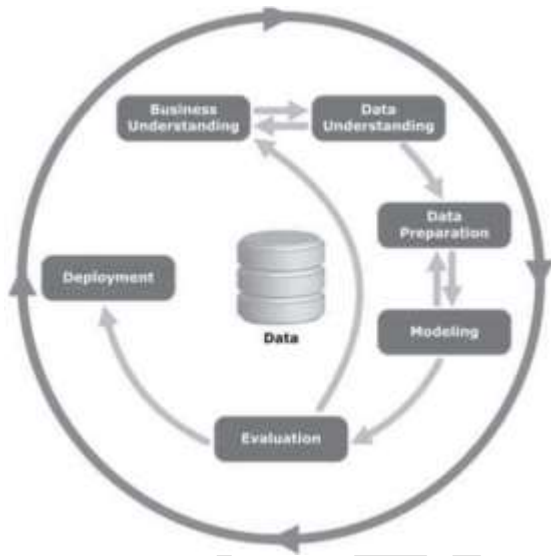


Fig 3: Data Flow Diagram

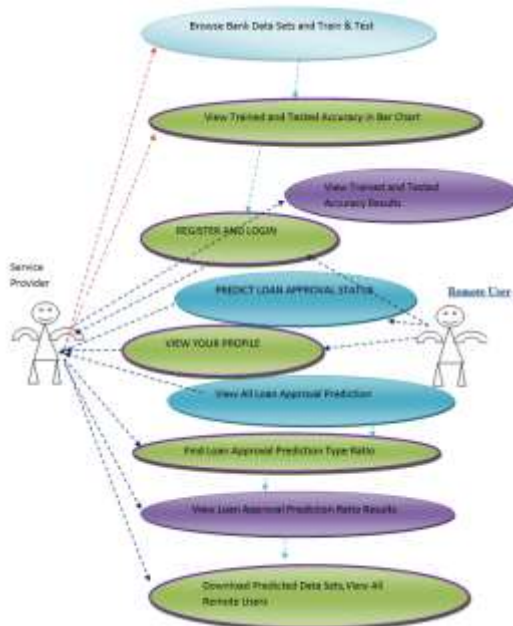


Fig 4 Use Case UML Diagrams

4. CONCLUSION

This study claims that both datasets' prediction accuracy is excellent. When a customer is experiencing a catastrophe, for example, the algorithm may not be able to forecast the right outcome. This study can determine if a prospective consumer would return a loan, and its accuracy is excellent. The most crucial variables for determining (if the customer would have been) are loan length, loan amount, age, and income. Zip code and credit history are the two most crucial variables in determining the loan applicant's category.

5. FUTURE SCOPE

Future research on the effectiveness of machine learning algorithms for loan approval prediction may also be conducted using thorough analyses of other machine learning methods than these.

6. REFERENCES

- [1] Amruta S. Aphale and R. Prof. Dr. Sandeep. R Shinde, "Predict Loan Approval in Banking System Machine Learning Approach for Cooperative Banks Loan Approval", International Journal of Engineering Trends and Applications (IJETA), vol. 9, issue 8, 2020
- [2] Loan Prediction Using Ensemble Technique, International Journal of Advanced Research in Computer and Communication Engineering Vol. 5, Issue 3, March 2016
- [3] Exploratory data analysis
- [4] Pandas Library
- [5] MeanDecreaseAccuracy <https://dinsdalelab.sdsu.edu/metag.stats/code/randomforest.html>
- [6] Kumar, S., Gunjan, V.K., Ansari, M.D., Pathak, R. (2022). Credit Card Fraud Detection Using Support Vector Machine. In: Gunjan, V.K., Zurada, J.M. (eds) Proceedings of the 2nd International Conference on Recent Trends in Machine Learning, IoT, Smart Cities and Applications. Lecture Notes in Networks and Systems, vol 237. Springer, Singapore. https://doi.org/10.1007/978-981-16-6407-6_3.