

## **LOAN APPROVAL PREDICTION**

**Abhishek k**

PG, Student

The Oxford College of Engineering,

Bommanahalli,

Bengaluru- 560068

[Abhishekkmca2025@gmail.com](mailto:Abhishekkmca2025@gmail.com)

**Manivannan jayachandran**

Associate Professor

Dept. of MCA

The Oxford College of Engineering,

Bommanahalli, Bengaluru- 560068

[manivananmca@gmail.com](mailto:manivananmca@gmail.com)

### **ABSTRACT**

Customer satisfaction and institutional profitability have always relied on the aspect of Credit approval process which is a core activity of financial institute. Conventional systems used to acquire loans are in most cases saddled with manual checks, rules based calculations, and subjectivity which makes the process slow, skewed and prone to errors. As financial technology evolves, smart, transparent solutions, which can more effectively estimate the creditworthiness of applicants, are becoming more and more necessary. In this study, a new predictive system in Credit approval process was presented by the researchers who combine Automated learning with deep learning (DL) to maximize precision and equity. The parameters used by the system include income level, working history, credit rating, debt to income ratio, and history of repayment. The system can train and test Algorithms like Random Forest and Support Vector Machines, XGBoost and Artificial Neural Network,

thus reaching very high accuracy rates in contrast to the established methods. In addition, the framework will add explainable AI techniques that aim to give applicants and the loan officers a clear picture of the basis of the prediction proving against prejudice and promoting credibility.

**KEYWORDS:** *Loan Approval Prediction, Credit Risk Assessment, Machine-Learning, Deep-Learning, Explainable-AI, Financial Technology*

### **INTRODUCTION**

Loan approval also has significant role to play in the financial industry and this directly affects the institutional development and client confidence. In conventional models, loan applications are assessed by credit officers who consider the aspects as the income, working-stability, borrowing record, and credit quality before assuming the answer. Although this manual approach was adopted decades ago, it is commonly slack, cost ineffective and open to human bias. As online banking has emerged and financial data can now be found online in large volumes, There is a growing demand for to have much more reporting, unbiased and efficient decision-

making systems. The recent achievements in artificial intelligence and, in particular, machine and deep learning provide the ability to predict repayment potential of an overall candidate with a greater precision. The automated process of risk assessment by including predictive models helps banks and lending agencies defaults, enhance customer service, and foster equity and as result, it helps establish modern inclusive banking practices.

## **LITERATURE SURVEY**

The measurement of credit risk on the research agenda because the default of the loans has a direct impact on the financial stability. Previously done research was too dependent on more interpretive tools like the logistic regression and decision trees which were not able to explain more complex non-linear interactions in the borrower data. Researchers started to use machine learning as the datasets started getting bigger and more varying. Random Forest and Gradient boosting Methods have risen in importance owing to their over-fitting mitigating and predictive scores; especially on lopsided data sets.

Several deep-learning approaches have been examined in recent years to reveal

the hidden behavioural trends in customer transactions and repayment record-keeping, including Network-based learning models networks. Together with predictive performance, the notions of fairness and transparency have turned into high-priority issues. Research also points to automated systems as potential sources of bias related to gender, age, or location, a fact that gave way to explainable AI systems such as SHAP and LIME. These frameworks do not only promote a better trust, but they also present coherence to financial institutions and applicants on the approval decisions. In spite of these improvements several existing approaches are inadequate to trade high predictive quality and interpretability, thus there remains room to improve hybrid initiatives.

## **EXISTING WORK**

The way the current financial institutions process borrowings requires improvement as it continues to rely heavily on the outdated manual judgments. A loan officer usually verifies the financial information provided by the applicant, record of employment, property records of the borrower, credit rating, and any other supplementary documents when making a decision. Though this is effective to an extent, it tends to be very time consuming and much manpower is required. As a loan file must be monitored in comparison with a number of eligibility parameters, assessment

becomes repetitive and subject to errors. The other weakness of the current system is that it depends on restricted historical data. Most banks pay attention merely to credit report or gross earnings of the account candidate and this may not give a perfect picture of financial conduct of the candidate. This will lead to the disadvantage of deserving candidates being rejected based on its strict policies, and some high-risk candidates can be allowed through. This unbalance exposes financial institutions to the risk of incurring non-performing assets.

Moreover, conventional method is not consistent, as even the officers can have different sources of understanding of eligibility rules. As the number of loans applications continues to increase on daily basis, the manual system prevails with challenges of accuracy, fairness and efficiency. That is why, an automated predictive solution is highly required.

## **PROPOSED SYSTEM**

The suggested system presents a machine-based and data-driven loan prediction model that optimally uses machine learning methods to justify or refute the application of an applicant with more accuracy, speed and objectivity. The system does not use pure judgment but evaluates extensive

numbers of features including applicant earnings, co-applicant earnings, education, loan sum, home type, married state, and above all, credit score. These are processed using algorithmic models that have been trained with historical data about loans, thus the model can understand trends that can be related to both successful and non successful loans.

The point would be to reduce the element of subjectivity, and that all applications would be evaluated against the same intelligent parameter. Classifiers in machine learning Among them are Logistic Regression, Random Forest, and Decision Tree, and Support Vector machine could be used to predict whether a loan should be rejected or approved. Correct performance measures will be used to confirm that the model is valid to include, accuracy, precision, recall, and ROC index.

Using a predictive system as a part of loan approval makes the workflow much faster and manpower requirements decrease by a major extent which can be passed on to the financial institution and its workforce. The system would offer prompt numbers that help a loan officer in the decision-making process and minimizing a scenario where human bias and dissimilarity can occur. It also enhances the customer experience through enhanced quicker response and increased transparency of the approval process. On the long run, the

suggested system can not only be advantageous to the applicants through being fair, but also protect the banks against any possible default by early detection of the high-risk borrowers.

## METHODOLOGY

The approach that should be implemented with the loan prediction project operates according to an organised strategy and Methodology initiated by acquiring of data and preprocessing, moves on to building of a model, training, testing and evaluation. The former involves the initial activity of data acquisition, during which records on historical loan application are collected. The details about the applicants, which are normally available in these datasets, include Gender, marital status, dependents, education, type of employment, income, amount of loan, credit history, and area of property, as well as final outcomes of the loan.

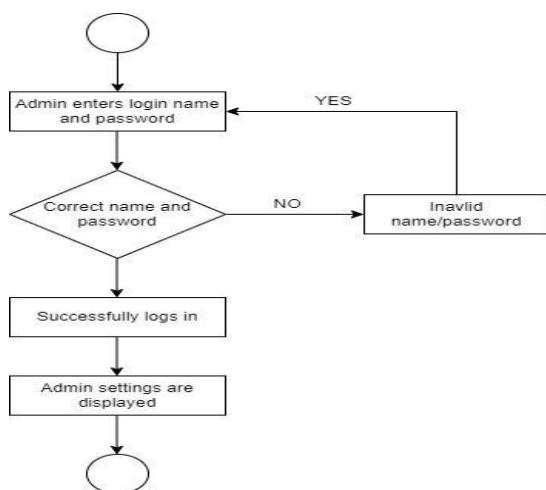


Figure 1 Activity flow

Determining the dataset is the first step followed by important activities of data preprocessing before analysis. This involves managing missing values, distributing factors that are categorical, and censoring numerical indicators in such a way that the algorithms could gain a perfect understanding of the information. At this point any reforming or useless attributes will be eliminated and the data will be separated into training and test sets, typically at 80:20 or 70:30 and allows fair comparison.

The second step is model building during Using several machine learning methodologies (including Approaches such as Logistic Regression, Decision Trees, Random Forest, Naïve Bayes, and SVMs machines) were used. Training of the models on the training dataset makes the models learn patterns that relationships between input features and loan approval outputs. When the training is over, the models are tested on the tested testing dataset and assess their performance.

Last but not least, the comparison of model performance is done with metrics Considering accuracy, precision, recall, and F1-score and ROC index. The winning model is then deployed. With such systematic approach, the project can have the reliability of the predictive system, its scalability and its relevant services That will enable financial institutions make evidence-based decisions.

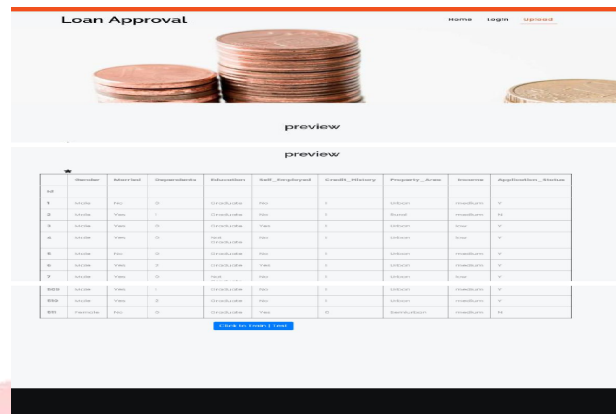
## EXPERIMENTAL RESULTS

The experimental stage of a loan prediction project involved the preparation of trainings and testings of different models of ML based on the prepared dataset. To make sure the models could be fairly evaluated, the datasets was considered in terms of trainings and testings in an 80:20 proportion. Applying its individual algorithm, Namely, Logistic Regression, Decision Tree, and Random Forest, Naïve Bayes, and Support Vector machine to training data and testing it on the undisclosed testing data set were incorporated.

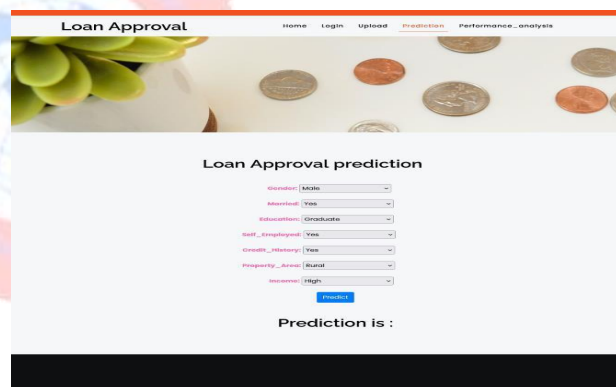
Accuracy was used to measure the performances of such models, but other measurements PM-precision, PM-recall, and PM-F1-score were also taken into account. Logistic Regression and Random Forest was more perfect than other models implying that these models would be effective in predicting loan eligibility. Decision Tree and Naïve Bayes have had fair performance with slight differences in accuracy with the K-Nearest Neighbors, which had low performance owing to the sensitivity to data distribution.

Given that, the experiments showed that

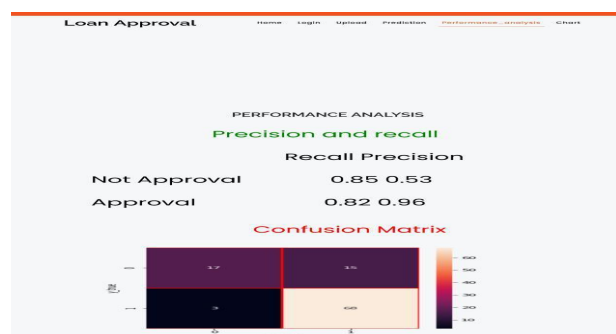
ML methods could be effective to detect the patterns of applicant information and determine the result of a loan application. Random Forest did the best generalization, therefore it could be offered as an appropriate candidate of real-world application.



*Figure 2 Upload page*



*Figure 3 Prediction page*



*Figure 4 Performance analysis page*

## CONCLUSION

The loan prediction project demonstrates how machine learning can be effectively applied to enhance the efficiency and accuracy of the loan approval process. Traditional manual evaluation methods are often time-consuming, inconsistent, and prone to human bias. By introducing predictive models, the decision-making process becomes faster, more reliable, and transparent. The experimental results showed that algorithms such as Logistic Regression and Random Forest provided strong performance, with Random Forest emerging as the most balanced and dependable model.

This system not only benefits financial institutions by minimizing the risk of defaults and reducing operational workload but also improves customer experience by offering quicker responses and fair assessments. Applicants are evaluated on multiple parameters, ensuring that the final decision is data-driven rather than subjective.

Overall, the proposed predictive system can serve as a valuable decision-support tool in banking and finance, helping institutions handle the increasing volume of applications more effectively. With further refinement, the system can be expanded for large-scale, real-world deployment.

## REFERENCES

- [1] Patel, R., & Desai, M. (2020). Loan Prediction System using Machine Learning Techniques. *International Journal of Engineering Research & Technology*, 9(7), 1152–1156.
- [2] Singh, A., & Kaur, P. (2021). A Comparative Study of Machine Learning Algorithms for Loan Approval Prediction. *International Journal of Computer Applications*, 183(23), 10–15.
- [3] Sharma, R., & Gupta, S. (2019). Predictive Analytics in Banking Sector Using Classification Models. *International Journal of Scientific & Technology Research*, 8(11), 2451–2456.
- [4] Kumar, V., & Chatterjee, A. (2022). Application of Data Mining in Financial Risk Management: Loan Default Prediction. *International Journal of Data Science and Analytics*, 14(2), 89–98.
- [5] Han, J., Kamber, M., & Pei, J. (2011). *Data Mining: Concepts and Techniques*. Elsevier, Morgan Kaufmann Publishers.