As per UGC guidelines an electronic bar code is provided to secure your paper

International Journal for Modern Trends in Science and Technology
Volume 12, Issue 04, pages 1330-1335
ISSN: 2455-3778 online
Available online at: http://www.ijmtst.com/vol12issue04.html
DOI: https://doi.org/10.5281/zenodo.19702121

Automated Loan Approval Decision Supportive System Using Predictive Analytics

L. Bujji Babu, G. Yaswanthisiri, Ch. Mythili, J. Padma Vamsi Babu

Department of Computer Science and Engineering, D.N.R. College of Engineering & Technology, Balusumudi, Bhimavaram, Andhra Pradesh, India

To Cite this Article
L. Bujji Babu, G. Yaswanthisiri, Ch. Mythili & J. Padma Vamsi Babu (2026). Automated Loan Approval Decision Supportive System Using Predictive Analytics. International Journal for Modern Trends in Science and Technology, 12(04), 1330-1335. https://doi.org/10.5281/zenodo.19702121

Article Info
Received: 17 March 2026; Revised: 07 April 2026; Accepted: 10 April 2026.

KEYWORDS

Loan Approval, Predictive Analytics, Machine Learning, Random Forest ,Logistic Regression, Decision Tree, Credit Risk, SVM, KNN, Financial Technology.

ABSTRACT

The rapid growth of the banking and financial sector has led to a significant increase in loan applications, making manual credit evaluation processes time-consuming, error-prone, and inefficient. Traditional loan approval methods rely heavily on human judgment and predefined rules, which often fail to capture complex patterns in applicant data and increase the risk of loan defaults. This paper proposes an Automated Loan Approval Decision Support System using Predictive Analytics to enhance accuracy, speed, and reliability in loan approval decisions. The system employs machine learning algorithms such as Logistic Regression, Decision Tree Classifier, Random Forest Classifier, K-Nearest Neighbors, and SVC to analyze applicant financial and demographic data. By leveraging historical loan records and predictive modeling techniques, the system evaluates creditworthiness and predicts loan approval outcomes effectively. The proposed solution improves decision consistency, reduces processing time, minimizes financial risk, and supports banks in making data-driven lending decisions.

1. INTRODUCTION

The concept of automated decision-making in financial services has evolved significantly with the advancement of data science, machine learning, and predictive analytics. The traditional loan approval process in financial institutions relies on manual verification, making it time-consuming, error-prone, and sometimes inconsistent. Issues such as delayed processing, human bias, and incorrect risk assessment can lead to financial losses or rejection of eligible applicants [1, 2, 7].

With advancements in predictive analytics and machine learning, loan approval systems can be automated to improve efficiency and accuracy. By analyzing historical data, algorithms like Logistic Regression and Decision Trees can predict an applicant's creditworthiness based on factors such as income, employment, and credit history [3, 4, 8].

The proposed Automated Loan Approval Decision Supportive System uses predictive analytics to provide fast, reliable, and data-driven decisions, reducing manual effort and improving overall risk management in financial institutions. The system evaluates multiple factors simultaneously, including income levels, credit history, loan amount, employment status, and property area, to arrive at an objective decision [5, 9].

1.1 Purpose

The Automated Loan Approval Decision Supportive System is designed to help financial institutions make accurate and quick loan decisions. It replaces traditional manual processes with a data-driven approach, reducing time and minimizing errors. The system analyzes applicant details such as income, credit history, and employment status, and uses machine learning techniques to predict loan eligibility. This improves consistency in decision-making, reduces human bias, and enhances risk assessment.

1.2 Motivation

In today's digital world, traditional loan approval systems are becoming inefficient due to their manual and time-consuming nature. These methods often lead to delays, errors, and inconsistent decisions because of human involvement. There is a strong need for faster and more accurate systems to evaluate loan applications. Financial institutions face risks when loans are wrongly approved or rejected. By using predictive analytics, this project provides a data-driven solution that improves efficiency, reduces risk, and ensures fair and consistent loan approval decisions.

2. LITERATURE SURVEY

Various studies show that traditional loan approval systems are slow, manual, and prone to errors and bias [6, 7, 8, 9, 10]. Researchers have applied machine learning algorithms such as Logistic Regression, Decision Trees, and Random Forest to predict loan eligibility based on applicant data like income, credit score, and employment status.

Li et al. [2] employed machine learning models to predict credit risk based on loan profitability using data from Chinese SMEs. The study demonstrated that loan profitability is a strong predictor of credit risk, improving the accuracy of risk assessment models. Kesraoui et al. [1] examined the relationship between credit and liquidity risk on banking margins using panel data from MENA countries, finding that both risks significantly affect bank margins.

Naili and Lahrichi [3] provided a comprehensive review of factors influencing banks' credit risk, including macroeconomic factors, bank-specific characteristics, and regulatory environments. Zhang and Yu [4] developed a system using machine learning to improve credit risk prediction, employing algorithms including Decision Trees, SVM, KNN, Random Forest, and XGBoost.

Abdelaziz et al. [5] focused on interactional relationships between credit risk, liquidity risk, and bank profitability in the MENA region, finding that both risks have negative impacts on bank profitability. These predictive models improve accuracy, speed, and consistency in decision-making. However, challenges such as data bias and lack of transparency still exist [6, 10]. Overall, the literature highlights that predictive analytics provides an efficient and reliable solution for modern loan approval systems.

3. PROPOSED METHODOLOGY

The proposed system introduces an automated decision support framework using predictive analytics and machine learning techniques. It processes historical loan data, performs data preprocessing and feature engineering, and applies multiple classification algorithms to predict loan approval outcomes. The system evaluates applicant creditworthiness in real time and provides accurate approval or rejection recommendations [4, 9].

3.1 System Architecture

The system architecture consists of three primary modules: Applicant Data Input, Machine Learning Model Pipeline, and Output Decision Engine. Applicant financial and demographic data is first collected and preprocessed, then passed through multiple trained classifiers. Risk assessment generates the final loan approval or rejection outcome.

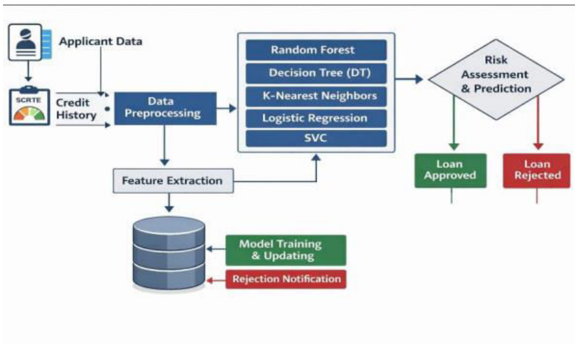

Fig. 1: System Architecture

3.2 Use Case Diagram

The use case diagram illustrates the interactions between the primary actors — the Applicant and the Bank Officer — and the core system functionalities including loan application submission, data verification, model-based prediction, and result viewing.

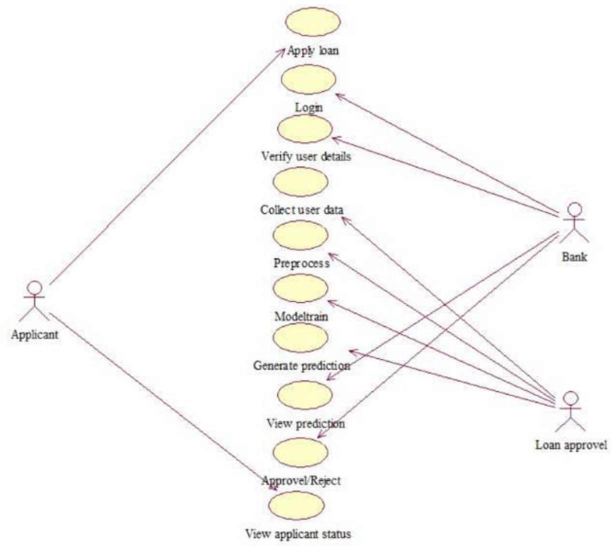

Fig. 2: Use Case Diagram

3.3 Class Diagram

The class diagram defines the structural relationships between the Applicant, BankOfficer, LoanApproval, and ML Model classes. The ML Model class serves as the base for Logistic Regression, Decision Tree Classifier, Random Forest Classifier, and KNeighbors Classifier sub-classes, each with their respective attributes and methods.

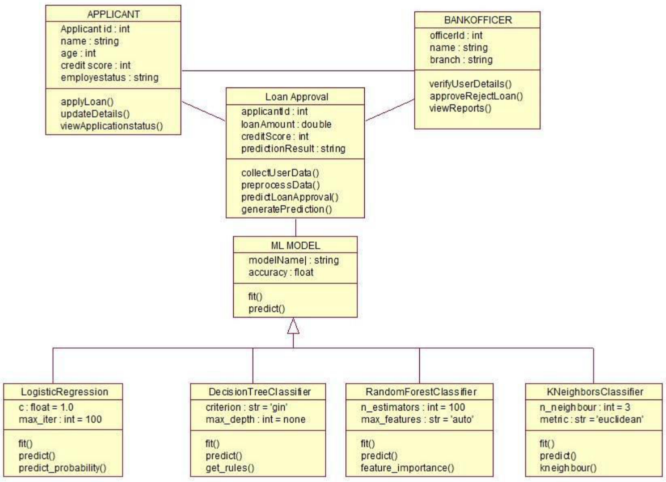

Fig. 3: Class Diagram

The sequence diagram below shows the time-ordered message flow between the Applicant, BankOfficer, and LoanApproval modules — from loan application submission through model training and final prediction.

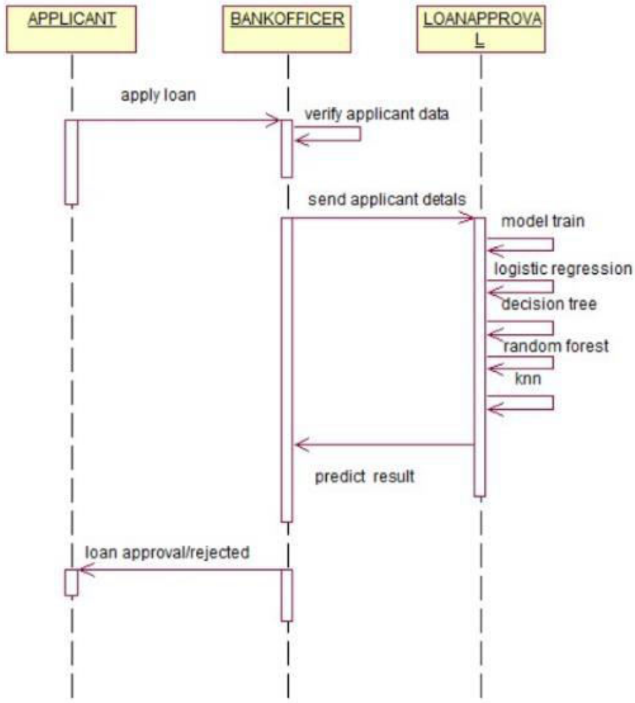

Fig. 4: Sequence Diagram

The activity diagram models the concurrent and sequential activities in the system workflow — from application detail collection through model training (Logistic Regression, Decision Tree, Random Forest, KNN) to the final loan approval or rejection decision.

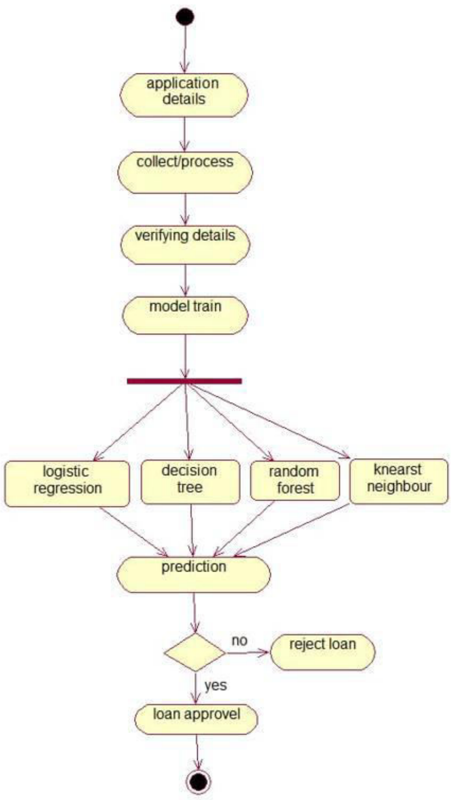

Fig. 5: Activity Diagram

3.4 Dataset

The dataset used is a standard loan eligibility dataset (Dataset.csv) containing historical loan application records with the following features: Loan_ID, Gender, Married, Dependents, Education, Self_Employed, ApplicantIncome, CoapplicantIncome, LoanAmount, Loan_Amount_Term, Credit_History, Property_Area, and Loan_Status (target variable).

Data preprocessing steps include: (1) Missing value imputation using median for numerical and mode for categorical features; (2) Feature engineering — Total_Income = ApplicantIncome + CoapplicantIncome, log transformations applied to income and loan amount; (3) Label Encoding of categorical variables; (4) Standard Scaling of feature values; (5) SMOTE oversampling to handle class imbalance [7].

3.5 Evaluation Metrics

The following evaluation metrics were used to compare model performance:

- Accuracy: Ratio of correctly predicted instances to total instances.
- Precision: Ratio of true positive predictions to total positive predictions.
- Recall: Ratio of true positives to actual positives (sensitivity).
- F1-Score: Harmonic mean of precision and recall.
- Confusion Matrix: Visual representation of true/false positives and negatives.

The dataset was split 75:25 for training and testing. SMOTE was applied only on training data to prevent data leakage.

4. RESULTS

Five machine learning models were trained and evaluated on the preprocessed loan dataset. The table below summarizes the accuracy and performance metrics of each model

Algorithm	Accuracy	Precision	F1-Score
Logistic Regression	68.72%	0.68	0.69
Decision Tree	73.93%	0.74	0.74
Random Forest	77.73%	0.78	0.78
KNN	68.72%	0.69	0.69
SVM	69.67%	0.70	0.70

Table 1: Comparative Model Performance

Random Forest achieved the highest overall accuracy of 77.73% with the best recall for approved loans (92%), making it the most effective model for identifying eligible applicants. The confusion matrix for Random Forest showed 75 true negatives, 89 true positives, 39 false positives, and 8 false negatives [2, 4].

HOME PAGE

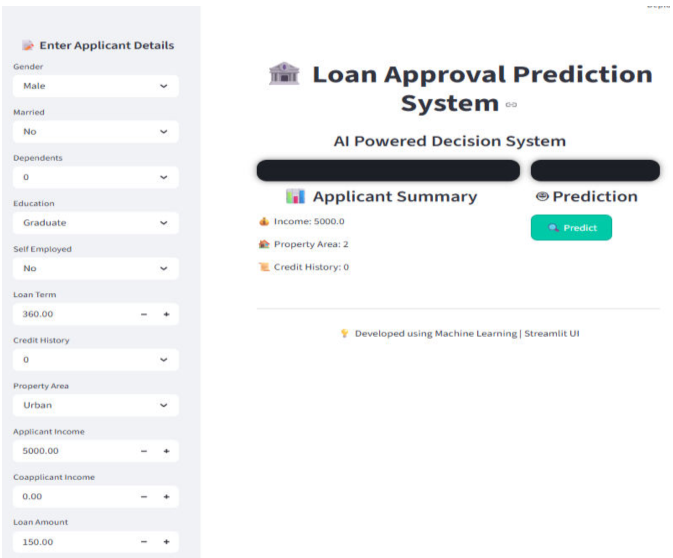

Fig. 7: Loan Approval Prediction System — Home Page

APPLICANT DETAILS

- **Bank officer enter details**

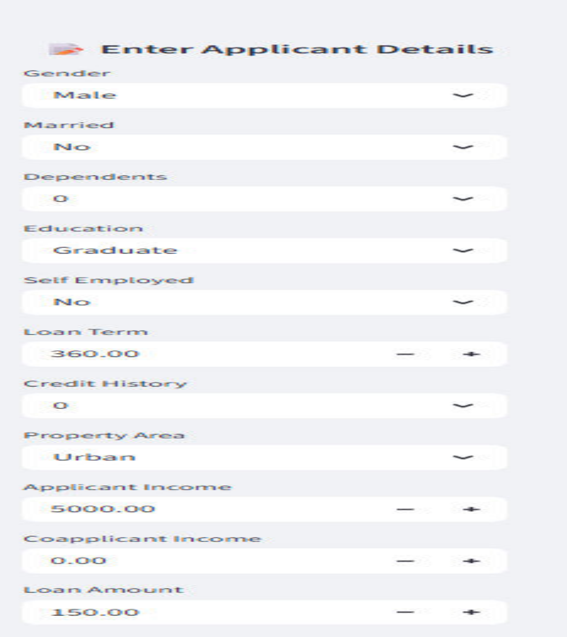

Fig. 8: Bank officier enter details

LOAN PREDICTION

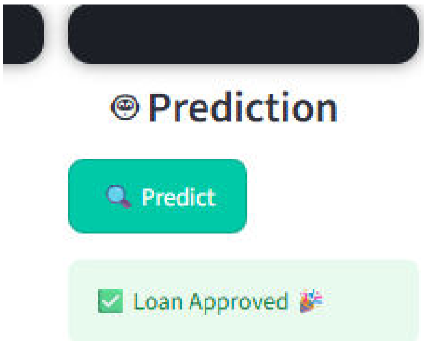

Fig. 9: Loan Approved Result Screen

LOAN APPROAVAL SCREEN

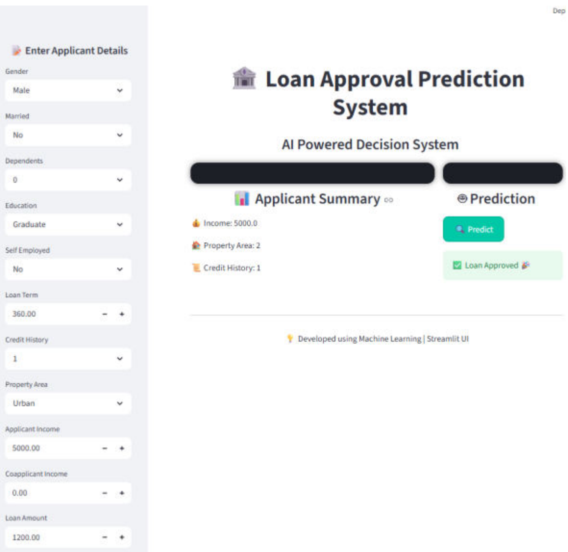

Fig. 10: Loan Not Approved Result Screen

5. CONCLUSION

This paper demonstrates the transformative potential of predictive modeling in enhancing the loan approval process within financial institutions. By comparing Logistic Regression, Decision Tree Classifier, Random Forest Classifier, K-Nearest Neighbors, and SVC, the study shows that Random Forest achieves the best overall accuracy (77.73%) and the highest recall for approved loans (92%), making it the most effective model for identifying eligible applicants.

Automated loan approval systems powered by AI-driven models reduce processing time, minimize human bias, and enhance financial institutions' ability to assess creditworthiness. The Streamlit-based deployment provides an accessible and user-friendly interface that can be readily adopted by banking professionals.

6. FUTURE SCOPE

- Integration with external credit bureaus and financial databases for real-time credit scores.
- Extension to multiple financial services such as credit card approval and insurance risk assessment.
- Implementation of advanced deep learning models for improved prediction accuracy.
- Upgrade to provide instant loan approval decisions using real-time data processing.
- Incorporating Explainable AI (XAI) to provide transparent reasons for approval or rejection.
- Cloud and mobile deployment for enhanced accessibility and scalability.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

REFERENCES

[1] Kesraoui, A.; Lachaab, M.; Omri, A. The impact of credit risk and liquidity risk on bank margins during economic fluctuations: Evidence from MENA countries with a dual banking system. Appl. Econ. 2022, 54, 4113–4130.

[2] Li, Z.; Liang, S.; Pan, X.; Pang, M. Credit risk prediction based on loan profit: Evidence from Chinese SMEs. Res. Int. Bus. Financ. 2024, 67, 102155.

[3] Naili, M.; Lahrichi, Y. The determinants of banks' credit risk: Review of the literature and future research agenda. Int. J. Financ. Econ. 2022, 27, 334–360.

[4] Zhang, X.; Yu, L. Consumer credit risk assessment: A review from the state-of-the-art classification algorithms, data traits, and learning methods. Expert Syst. Appl. 2024, 237, 121484.

[5] Abdelaziz, H.; Rim, B.; Helmi, H. The interactional relationships between credit risk, liquidity risk and bank profitability in MENA region. Glob. Bus. Rev. 2022, 23, 561–583.

[6] Huang, Y. et al. BigTech credit risk assessment for SMEs. China Econ. Rev. 2023, 81, 102016.

[7] Bhatore, S.; Mohan, L.; Reddy, Y.R. Machine learning techniques for credit risk evaluation: A systematic literature review. J. Bank. Financ. Technol. 2020, 4, 111–138.

[8] Singh, V. et al. Prediction of modernized loan approval system based on machine learning approach. In Proc. CONIT 2021, Hubli, India, pp. 1–4.

[9] Weng, C.; Huang, C. A hybrid machine learning model for credit approval. Appl. Artif. Intell. 2021, 35, 1439–1465.

[10] Simos, T.E. et al. A multi-input with multi-function activated weights and structure determination neuronet for classification problems. Appl. Soft Comput. 2022, 127, 109351.