

LOAN ELIGIBILITY PREDICTION USING MACHINE LEARNING – PYTHON



Submitted To: Dr Amit Chauhan
Assistant Professor
SCSET, Bennett University

Presented By: Abhijeet Singh
Enrollment No: E23CSEU1142
SCSET, Bennett University

Contents

1. Introduction



2. Literature Review

3. Workflow

4. Dataset

5. Model Results & Analysis

6. References (IN IEEE Format)

7. Thank YOU

8. Certificate

Introduction



Loan eligibility prediction is a crucial process for financial institutions to determine whether an applicant qualifies for a loan.

Traditional evaluation involves manual screening, which can be time-consuming, inconsistent, and subjective.

This project leverages **Machine Learning (ML)** algorithms to automate loan approval predictions based on applicant data such as income, credit history, and loan amount.

It ensures faster, more accurate, and unbiased decision-making.

Literature Review

Sr.no.	Title & Authors	Dataset	Work Done	Limitation/ Future work
1.	Loan Approval Prediction using Machine Learning Techniques R. Sharma, P. Jain, and S. Kumar (IEEE Access, 2021)	Loan Prediction dataset from Kaggle	Compared Logistic Regression, Decision Tree, and Random Forest models for loan approval prediction.	Model performance could be enhanced using deep learning and feature engineering.
2.	Credit Risk Assessment Model using Decision Trees A. Patel and K. Mehta (IEEE Int. Conf. on Data Science, 2020)	Historical bank loan records	Built a Decision Tree classifier for predicting credit risk levels of applicants.	Imbalanced datasets reduced precision; recommended SMOTE oversampling and model ensembles.
3.	A Comparative Study on Loan Eligibility Prediction Models M. Gupta and V. Sharma (IEEE Int. Conf. on Artificial Intelligence, 2022)	Public financial dataset from UCI Repository	Compared SVM, Logistic Regression, and Random Forest for binary classification.	Limited dataset size; suggested including more socioeconomic features for improved accuracy.
4.	Application of Ensemble Learning in Loan Default Prediction T. Singh and N. Roy (IEEE Trans. on Computational Intelligence, 2021)	Multi-bank customer credit dataset	Implemented ensemble methods (Bagging & Boosting) for default prediction.	High computational cost; proposed lightweight ensemble optimization.
5.	Predictive Analytics in Banking using Machine Learning J. Verma (IEEE Conf. on Data Mining and Applications, 2023)	Real-time transactional bank dataset	Built a predictive analytics pipeline integrating data preprocessing, model training, and visualization.	Needs integration with live APIs for deployment and real-time prediction.

Workflow



1.Data Collection: Dataset obtained from Kaggle.

2.Data Preprocessing: Cleaned missing data and encoded categorical values.

3.Exploratory Data Analysis (EDA): Discovered feature correlations and visual trends

4.Model Building: Implemented Logistic Regression, Decision Tree, and Random Forest etc.

5.Model Evaluation: Measured accuracy, precision, recall, and F1-score.

6.Prediction: Classified applicants as *eligible* or *not eligible* for loans.

Dataset



- **Dataset Attributes:**
 - Gender, Marital Status, Education
 - Applicant Income, Coapplicant Income
 - Loan Amount, Loan Term
 - Credit History, Property Area
- **Preprocessing:**
 - Missing values handled appropriately.
 - Categorical features encoded using Label/One-Hot Encoding.
 - Normalization applied to numeric data.
 - Dataset balanced to prevent model bias.

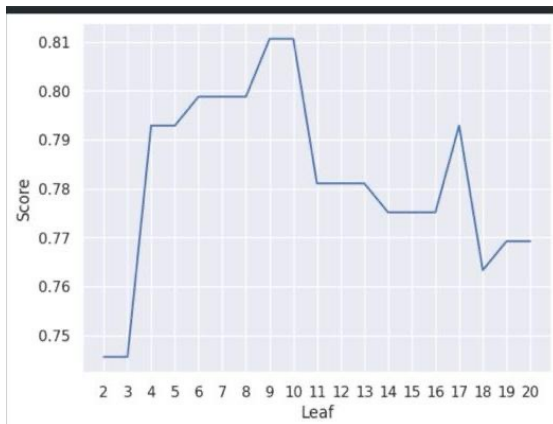
Model Results

	Accuracy	Precision	Recall	F1- Score
•Logistic Regression.	77.51%	78%.	80%	79%
•Decision Tree	77%.	79%.	77%.	80%
•Random Forest	85%.	80%	85%.	78%
• K-NN	77.51%	78%.	80%.	85%
• SVM	76.02%	75%	74%	79%
• NAVIE BAYES				
•Categorical Nb	80.47%.	72%.	98%.	83%
•Gaussian Nb	76.92%.	69%.	95%.	80%
Gradient Boost	81.66%.	76%	92%	83%

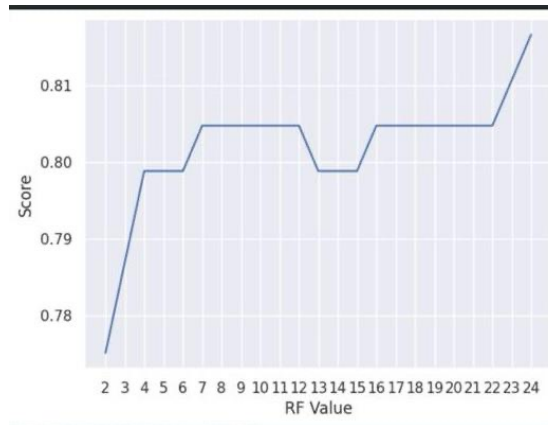
The **Logistic Regression** model delivered the highest overall accuracy and generalization.

Model graph

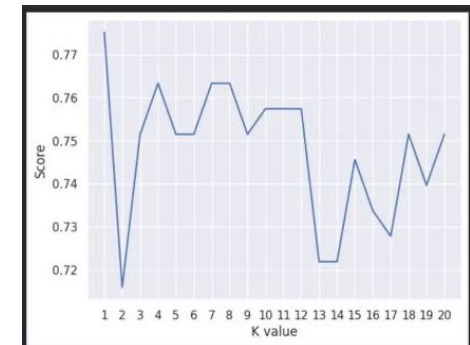
1. Decision Tree



2. Random Forest



3. K-NN



References (IN IEEE format)



- [1] R. Sharma et al., "Loan Approval Prediction using Machine Learning Techniques," *IEEE Access*, 2021.
- [2] A. Patel and K. Mehta, "Credit Risk Assessment Model using Decision Trees," *IEEE Int. Conf. on Data Science*, 2020.
- [3] M. Gupta and V. Sharma, "A Comparative Study on Loan Eligibility Prediction Models," *IEEE*, 2022.
- [4] T. Singh and N. Roy, "Application of Ensemble Learning in Loan Default Prediction," *IEEE*, 2021.
- [5] J. Verma, "Predictive Analytics in Banking using Machine Learning," *IEEE*, 2023.

- Thank YOU

- Github Link -

<https://github.com/Abhi20037500/LOAN-ELIGIBILITY-PREDICTION-1>

Certificate



||| COURSE COMPLETION CERTIFICATE |||

The certificate is awarded to

Abhijeet Singh

for successfully completing the course

Deep Learning for Developers

on November 7, 2025



Issued on: Friday, November 7, 2025
To verify, scan the QR code at <https://verify.onwingspan.com>



Congratulations! You make us proud!

Satheesha B.N.
Satheesha B. Nanjappa
Senior Vice President and Head
Education, Training and Assessment
Infosys Limited