IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

PREDICTION OF MODERNIZED LOAN APPROVAL SYSTEM BASED ON MACHINE LEARNING APPROACH

Authors: J. Leena Lahari, A. Jasmine, J. Banny, S. Jyothi, P Chaitanya Kumar
Under the Guidance of: Mrs. V Sravani Rajya Lakshmi, Assistant Professor
Department of Computer Science and Engineering
Visakha Institute of Engineering and Technology, Visakhapatnam

1. ABSTRACT

Technology has significantly enhanced human life, offering innovative solutions across various sectors, including banking. In the context of loan approval, machine learning (ML) plays a pivotal role by analyzing historical data to predict an applicant's eligibility. Due to limited financial resources and a high volume of loan applications, accurate predictions using classification algorithms—such as Logistic Regression, Random Forest, and Support Vector Machines—are crucial.

These models help minimize risk by identifying applicants likely to repay loans, thereby positively impacting the bank's profit and loss margins.

Machine learning not only accelerates the decision-making process but also ensures consistency and reduces human error. It allows banks to prioritize applicants with higher repayment potential, improving recovery rates and enhancing operational efficiency. The primary objective is to develop a robust, data-driven system to facilitate fair, accurate, and timely loan approvals.

2. KEYWORDS

Machine Learning, Loan Approval Prediction, Support Vector Machine(SVM), Classification Algorithms, Logistic Regression, Random Forest,Naïve Bayes, Data-driven Decision Making, Financial Risk Assessment,
Non-Performing Assets (NPA), Model Accuracy, Artificial Intelligence(AI), Banking Automation, Loan Eligibility, Predictive Modeling

3. INTRODUCTION

With the rapid digitization of the banking sector, the volume of online loan applications has increased dramatically. Manually evaluating these applications is not only time-consuming but also prone to human error. Artificial Intelligence (AI) offers an effective solution by enabling data-driven decision-making and automating the analysis of vast datasets.

An AI-based loan prediction system assesses applicant eligibility using features such as gender, marital status, number of dependents, education level, employment type, income, credit history, and property area. This enhances the bank's ability to make efficient, consistent, and fair decisions, thereby improving workflow efficiency, reducing turnaround time, and increasing customer satisfaction.

4. LITERATURE REVIEW

The literature survey plays a crucial role in laying the foundation for informed decision-making and effective planning. It involves examining time constraints, budgets, available resources, and the organization's technical capabilities. The survey helps determine project feasibility and identifies suitable operating systems, programming languages, development tools, and software components.

Moreover, reviewing existing systems, past research, and expert guidance ensures developers avoid common pitfalls and adopt best practices. An effective literature review reduces project risks, improves design accuracy, and sets a clear path for successful development and implementation.

5. METHODOLOGY

The study follows a structured approach to determine whether granting a loan to a particular individual is a safe financial decision. The project is divided into four main phases:

Data Collection: Gathering historical data on loan applicants.

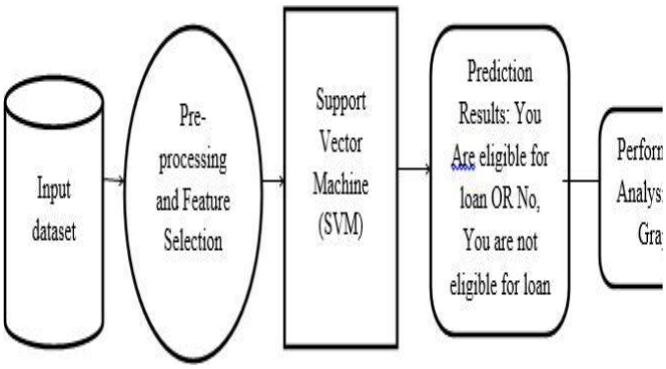
Model Comparison: Evaluating different machine learning models.

Model Training: Training the most effective algorithm.

Testing: Assessing the model's accuracy and performance.

Various algorithms such as Naïve Bayes, Support Vector Machines (SVM), Decision Trees, Logistic Regression, Random Forest, and Gradient Boosting were explored.
Among these, Naïve Bayes and SVM demonstrated superior performance in terms of speed and accuracy. The objective is to create a reliable, data-driven system to enhance loan approval efficiency and reduce financial losses due to non-performing assets (NPAs).

5. SYSTEM ARCHITECTURE



6. SYSTEM REQUIREMENTS

A. Hardware Requirements

- **Processor**: Pentium Dual Core
- **Hard Disk**: 120 GB
- **RAM**: 4 GB
- **Monitor**: 15" LED
- **Input Devices**: Keyboard, Mouse

B. Software Requirements

- **Operating System**: Windows 7/10
- **Programming Language**: Python
- **Libraries**: Scikit-learn, Pandas, NumPy, Matplotlib

7. ANALYZE AND PREDICTION In the actual dataset, we chose only 7 features:

•Gender:Male/Female
•Married: Applicant married (Y/N)
•Education: Applicant Education (Graduate/ Under Graduate) •Self_Employed: Self employed (Y/N)
• Credit_History: Credit history meetsguidelines
• Property_Area: Urban/ Semi Urban/Rural
- Income: Applicant income

• Application_Status: (Target) Loan approved (Y/N)

9. SYSTEM TESTING

Testing ensures the system's reliability and conformance to both functional and user expectations.

Various testing levels were conducted:

Unit Testing: Verifies the logic of individual functions or modules.

- **Functional Testing**: Validates the system's adherence to specifications.
- **System Testing**: Ensures all components work together as a cohesive unit.
- **White Box Testing**: Focuses on internal code structures.

Black Box Testing: Tests functionality without examining code.

Regression Testing: Ensures new updates don't break existing features.

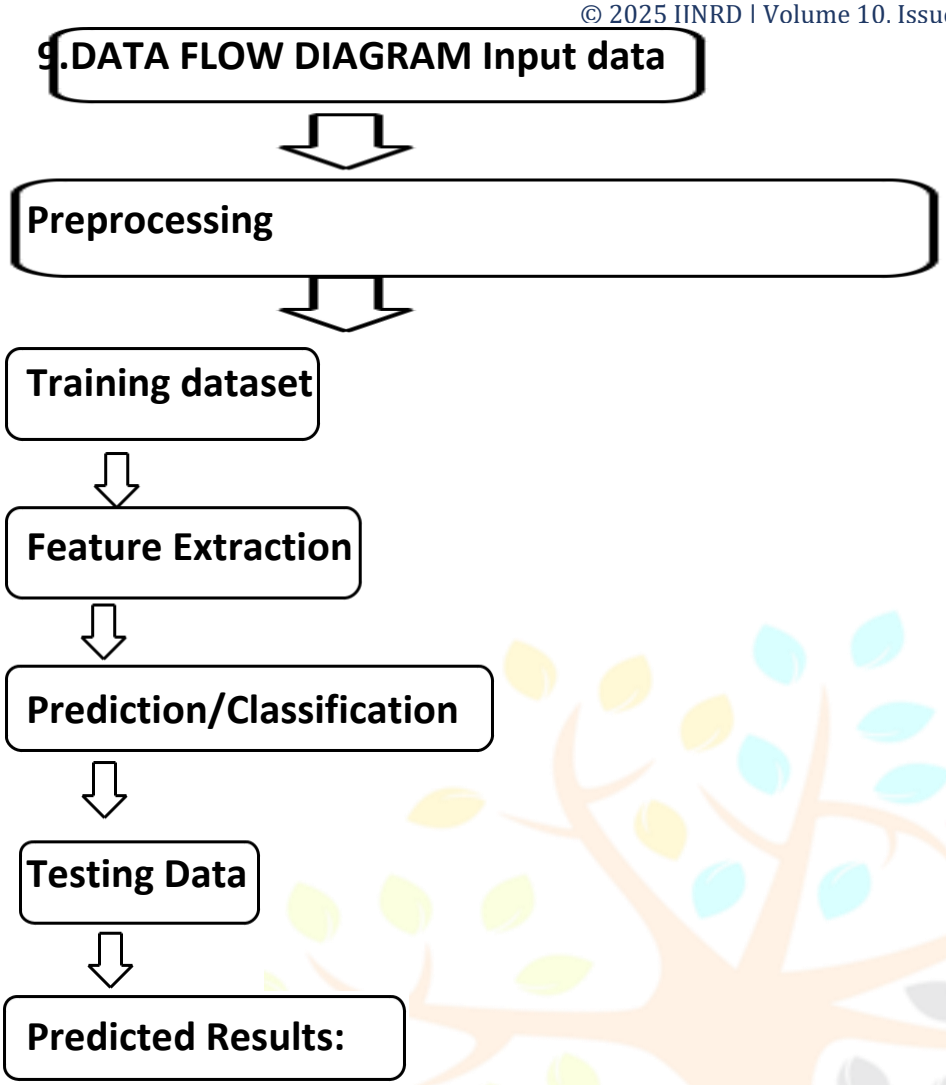
Performance Testing: Evaluates behavior under load.

These tests confirmed that the system processes input correctly, produces accurate predictions, maintains data integrity, and performs reliably under various conditions.

10. CLASS DIAGRAM(Tabular format)

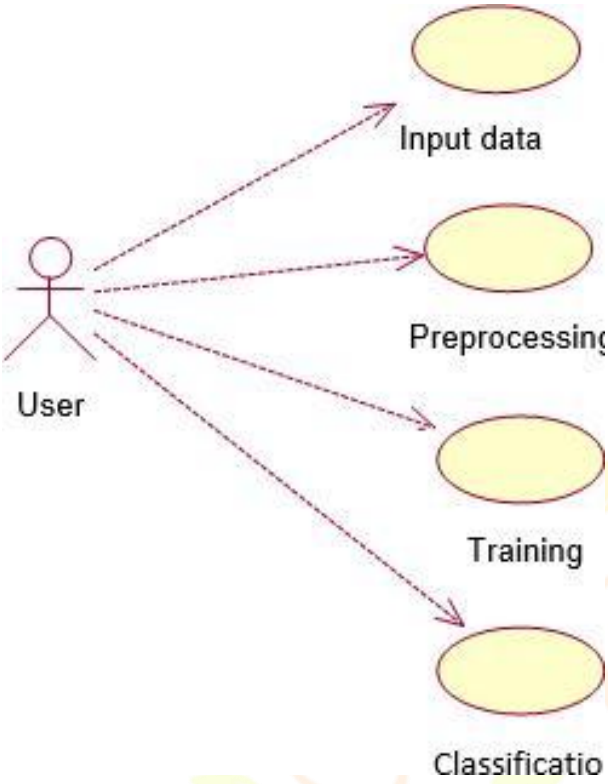
Component	Details
Input	- Input Data - Preprocessing
Process	- Feature Extraction - Classification
Output	- Final Classification - Display Result: • You are eligible for loan • You are not eligible for loan

9.DATA FLOW DIAGRAM Input data



You are eligible for loan OR No, You are not eligible for loan.

USE CASE DIAGRAM



8. Working Mechanism

The Support Vector Machine (SVM) algorithm is a powerful classification technique that identifies the optimal hyperplane for separating data points into different classes. In simple scenarios, SVM selects the hyperplane with the maximum margin—distance between the hyperplane and the nearest points (support vectors).

When applied to real-world data such as loan prediction, SVM effectively handles high-dimensional and non-linear data using the kernel trick, which transforms input features into a higher-dimensional space. This enables linear separation and improves classification accuracy. SVM's robustness to outliers and ability to generalize well make it suitable for complex tasks such as financial decision-making.

Stage	Description
Data Collection	Collecting historical data on loan applications, including features like income, credit history, and employment status.
Preprocessing	Handling missing

	values, encoding categorical data, and normalizing numerical values.
Feature Extraction	Selecting relevant features for better model performance and eliminating irrelevant or redundant data.
Model Selection	Comparing different machine learning models (e.g., SVM, Naïve Bayes) based on accuracy, speed, and robustness.
Training	Feeding the training data to the selected model to learn patterns and make decisions.
Testing	Evaluating model performance using test data to ensure accuracy and reliability.
Classification Output	Predicting whether a loan should be approved or not, based on the trained model.
Display Result	Communicating the final decision: "You are eligible for loan" or "You are not eligible for loan."

9. CONCLUSION

This research explores the application of Machine Learning— particularly the Support Vector Machine (SVM) algorithm—for predicting loan approval. The model achieved a training accuracy of 80% and a test accuracy of 82%, outperforming traditional manual approaches.

While the system performs well under standard conditions, it may face limitations when handling rare or unexpected scenarios, such as natural disasters or sudden financial downturns not reflected in the training data. Despite this, the model significantly enhances the loan approval process, offering consistent and data-driven decision support.

Future Work

Future studies can explore the integration of more advanced ML techniques such as XGBoost, LightGBM, and deep learning models to improve prediction accuracy further. Incorporating real-time data and socio-economic factors can also make the system more robust and adaptable to dynamic financial environments.

---

10. REFERENCES

1. Supriya, P., Pavani, M., Saisushma, N., & Kumari, V. (2019). *Loan Prediction by using Machine Learning Models*. International Journal of Engineering and Techniques, 5(2).

2. Madane, N., & Nanda, S. (2019). *Loan Prediction using Decision Tree*. Journal of the Gujarat Research History, 21(14s).

3. Vimala, S., & Sharmili, K.C. (2018). *Prediction of Loan Risk using NB and SVM*. International Conference on Advancements in Computing Technologies, 4(2), 110-113.

4. Hamid, A.J., & Ahmed, T.M. (2016). *Developing Prediction Model of Loan Risk in Banks using Data Mining*. Machine Learning and Applications: An International Journal (MLAIJ), 3(1), 1-9.

5. Arun, K., Garg, I., & Kaur, S. (2016). *Loan Approval Prediction based on Machine Learning Approach*. IOSR JournalofComputer Engineering (IOSR-JCE), 18(3), 79-81.