International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 2, Mar-Apr-2026, ISSN (Online): 2395-566X

An Intelligent Credit Risk Prediction Framework Using Machine Learning Algorithms

Ms.G.Naga Rani[1], Mangipudi V N S Sekhar Sarma[2], Khandavalli V V Lakshmi Srirama Karthik[3], Malla Karthik[4], Pabbineedi Vanshika[5], Ventru Hemanth Kumar[6]
[1]Assistant Professor, [2, 3, 4, 5, 6]B.tech Students Department of CSE, Pragati Engineering College, Surampalem, Andhra Pradesh, India

Abstract- The banking sector plays a vital role in the global financial system by providing loans to individuals and businesses for various purposes. While loans generate significant revenue through interest, there is always a risk that borrowers may fail to repay the loan, resulting in financial losses for lending institutions. Therefore, accurately predicting the risk level associated with a loan application is an important task for banks and financial organizations. Traditional loan approval processes rely heavily on manual analysis of customer information, which can be time-consuming and prone to human bias. With the advancement of machine learning techniques, automated systems can now analyse large amounts of financial data to support more efficient and accurate loan approval decisions. This study proposes a machine learning-based loan risk prediction system that analyses customer personal and financial attributes to determine the likelihood of loan default. The dataset used for this study contains multiple features commonly included in loan applications, such as credit history, checking account status, loan amount, employment status, and age of the applicant. Data preprocessing techniques including outlier removal, categorical encoding, and feature scaling are applied to prepare the dataset for model training. Several machine learning algorithms are implemented and compared, including Decision Tree, Random Forest, Support Vector Machine (SVM), Multi-Layer Perceptron (MLP), Naive Bayes, and a Stacking Ensemble model. The models are evaluated using performance metrics such as accuracy, precision, recall, F1-score, and confusion matrix analysis. Experimental results demonstrate that ensemble-based approaches provide improved predictive performance compared to individual machine learning models. The proposed system can assist financial institutions in making faster and more reliable loan approval decisions by identifying high-risk applicants before granting loans. By leveraging machine learning techniques, the system enhances the efficiency of credit risk assessment and supports more effective financial decision-making in the banking industry.

Keywords – Machine Learning, Credit Risk Prediction, Loan Approval System, Financial Data Analysis, Random Forest, Support Vector Machine, Ensemble Learning, Banking Analytics.

I. INTRODUCTION

Loans represent a fundamental component of the global financial system, enabling individuals and organizations to access financial resources for various purposes such as business expansion, education, housing, and personal development. Financial institutions, including banks and credit agencies, rely heavily on lending activities as a primary source of revenue through the interest generated from loan repayments. However, loan issuance also involves significant financial risks, particularly when borrowers fail to repay the borrowed amount within the agreed time frame. Such situations, commonly referred to as loan defaults, can result in substantial financial losses for lending institutions. Therefore, accurately assessing the creditworthiness and repayment capability of loan applicants is a critical task for financial organizations in order to minimize potential risks and maintain financial stability [3], [4].

Traditionally, loan approval decisions have been performed through manual evaluation processes in which bank officials analyse a customer's financial and personal information before granting credit. These assessments typically consider factors such as income level, credit history, employment status, loan amount, and repayment duration, commonly referred to as the "Five Cs of Credit," which include character, capacity, capital, collateral, and conditions [11]. Although this manual evaluation approach can provide useful insights, it is often time-consuming, susceptible to human bias, and difficult to apply consistently when processing large volumes of loan applications. With the rapid growth of digital banking services and financial technologies, there is an increasing demand for automated systems capable of efficiently evaluating loan applications and supporting data-driven decision-making processes.

Recent advancements in data analytics and artificial intelligence have enabled the application of machine learning techniques for financial risk assessment and credit scoring. Machine learning algorithms can analyse large-scale datasets containing customer financial attributes and identify hidden patterns that indicate the likelihood of loan repayment or default. By learning from historical customer data, these models can generate predictive insights that assist financial institutions in evaluating loan applications more accurately and efficiently [1], [2].

Several machine learning models have been successfully applied to credit risk prediction and loan approval systems. Algorithms such as Decision Trees, Random Forest, Support Vector Machines (SVM), Naïve Bayes, and Neural Networks have demonstrated strong capabilities in identifying high-risk loan applicants and predicting loan approval outcomes [3], [6], [9]. Among these methods, ensemble learning techniques, which combine multiple models to improve predictive performance, have gained significant attention in recent research. Ensemble-based approaches such as voting classifiers, boosting methods, and gradient boosting frameworks can improve prediction accuracy and reduce classification errors in complex financial datasets [8], [10].

In addition, advanced machine learning algorithms such as CatBoost and other boosting-based methods have been explored to enhance classification performance and handle large-scale financial datasets effectively. Comparative studies have demonstrated that boosting-based algorithms often outperform traditional classifiers due to their ability to capture nonlinear relationships and improve model generalization [7]. Furthermore, the availability of publicly accessible datasets, such as credit risk datasets from financial institutions and open data platforms, has facilitated the development and evaluation of machine learning-based credit risk prediction systems [12].

Motivated by these advancements, this study proposes a machine learning-based loan risk prediction system designed to analyse customer financial data and classify loan applicants as either low-risk or high-risk borrowers. The proposed system evaluates multiple machine learning algorithms and compares their predictive performance to determine the most effective model for credit risk prediction. Additionally, the analysis aims to identify the most influential financial attributes that contribute to loan approval decisions.

The primary objective of this research is to develop an intelligent predictive system that assists financial institutions in making faster, more accurate, and data-driven loan approval decisions. By integrating machine learning techniques into the loan evaluation process, the proposed system aims to reduce the risk of loan defaults while improving the overall efficiency and reliability of financial decision-making systems. The adoption of artificial intelligence in financial services also contributes to improving fairness and transparency in loan approval processes by reducing human bias and enabling consistent evaluation of applicants [13].

The remainder of this paper is organized as follows. Section II presents a review of existing research related to loan risk prediction and machine learning-based credit evaluation systems. Section III describes the system analysis and methodology used in this study. Section IV explains the implementation and model training process. Section V discusses the experimental results and performance evaluation of the proposed models. Finally, Section VI concludes the study and highlights potential directions for future research.

II. LITERATURE SURVEY

Credit risk prediction has become a significant research area in the financial sector due to the increasing demand for efficient and reliable loan approval systems. Financial institutions must carefully evaluate loan applications to minimize the risk of borrower default while maintaining profitable lending operations. Traditionally, credit risk assessment was performed manually by financial experts using statistical analysis and predefined decision rules. Although these traditional methods provided useful insights, they often struggled to process large volumes of financial data and capture complex relationships among multiple customer attributes [6].

With the advancement of artificial intelligence and data analytics, machine learning techniques have increasingly been applied to credit risk prediction systems. Machine learning algorithms can analyse large financial datasets and identify hidden patterns related to loan repayment behaviour. These models enable financial institutions to make data-driven decisions regarding loan approval and credit risk evaluation. Recent research has shown that machine learning models can significantly improve prediction accuracy compared to conventional statistical approaches when applied to financial datasets containing customer attributes such as income level, credit history, and employment status [1], [3].

Several studies have explored the application of regression-based models for predicting loan eligibility and credit risk. Regression techniques analyse relationships between dependent and independent variables to estimate the probability of loan approval. While regression models can provide reasonable prediction accuracy in certain financial datasets, they often struggle to capture nonlinear relationships

International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 2, Mar-Apr-2026, ISSN (Online): 2395-566X

between variables, which may limit their effectiveness in complex credit risk scenarios [1].

Classification-based machine learning algorithms have also been widely applied to loan approval prediction systems. Decision Trees and Random Forest models are commonly used techniques because they can handle both numerical and categorical variables while producing interpretable decision rules. In particular, Random Forest algorithms combine multiple decision trees to improve prediction performance and reduce overfitting. Research has demonstrated that Random Forest models can effectively analyse financial attributes and improve the accuracy of loan approval predictions [2], [4].

Support Vector Machine (SVM) models have also been extensively applied in credit risk prediction tasks. SVM algorithms work by constructing optimal decision boundaries that separate different classes within a dataset. When applied to loan approval problems, SVM models can effectively distinguish between low-risk and high-risk loan applicants based on their financial characteristics. However, SVM models often require careful parameter tuning and higher computational resources to achieve optimal classification performance [5].

Artificial Neural Networks (ANN) have gained increasing popularity in financial risk prediction due to their ability to model nonlinear relationships between variables. Neural networks consist of multiple layers of interconnected neurons that learn complex patterns from training data. Multi-layer perceptron (MLP) models, in particular, have been widely applied in credit scoring systems and financial forecasting applications because of their ability to capture hidden patterns in large datasets [6].

More recently, ensemble learning techniques have been introduced to further improve the performance of credit risk prediction models. Ensemble methods combine the outputs of multiple machine learning models to produce more robust and reliable predictions. Techniques such as voting classifiers, boosting algorithms, and stacking models have demonstrated improved predictive performance by leveraging the strengths of different classifiers. Studies have shown that ensemble-based systems can outperform individual machine learning algorithms in loan approval prediction tasks [8], [10].

In addition, advanced boosting algorithms such as CatBoost have been proposed to enhance classification performance and handle complex financial datasets effectively. Comparative studies indicate that boosting-based models often provide better accuracy and generalization capability compared to traditional classification techniques [7].

Despite the significant progress made in machine learning-based credit risk prediction systems, several challenges still remain. Many models are trained using limited datasets that may not fully represent real-world financial conditions. Furthermore, biased training data may lead to unfair or discriminatory loan approval decisions if sensitive attributes such as gender, ethnicity, or socio-economic status influence model predictions. Ensuring fairness and transparency in AI-based financial decision systems has therefore become an important research topic in recent years [13].

To address these challenges, the present study develops and evaluates multiple machine learning models for predicting the risk level of loan applicants based on their personal and financial attributes. The models are compared using several performance evaluation metrics to determine the most effective approach for credit risk prediction. By analysing different machine learning algorithms and ensemble techniques, this research aims to contribute to the development of more accurate, reliable, and efficient loan risk prediction systems for modern financial institutions.

III. SYSTEM ANALYSIS

Existing System

Traditional loan approval systems primarily rely on manual evaluation processes conducted by financial experts and bank officials. In these systems, the decision to approve or reject a loan application is based on the analysis of various customer attributes, including income level, credit history, employment status, loan amount, and repayment duration. These factors are typically assessed using predefined decision rules and statistical methods to determine the creditworthiness of an applicant. Financial institutions often consider the widely recognized Five Cs of Credit, which include character, capacity, capital, collateral, and conditions, as part of the evaluation process for loan approval decisions [11].

Although manual loan assessment methods have been widely used for many years, they exhibit several limitations when applied to modern financial environments characterized by large volumes of loan applications and complex financial datasets. Evaluating a large number of loan applications manually can be time-consuming and requires substantial human resources. Moreover, manual decision-making processes may introduce human bias and inconsistencies, which can negatively affect the fairness and reliability of loan approval outcomes [13].

To support the decision-making process, some financial institutions have adopted traditional statistical models such as logistic regression and rule-based decision systems. These models analyse historical financial data to estimate the probability of loan repayment and assist in evaluating the risk associated with each loan applicant. Regression-based models and basic classification approaches have demonstrated moderate effectiveness in credit risk analysis when applied to financial datasets containing customer attributes such as income, employment status, and credit history [1], [3].

However, traditional statistical approaches often struggle to capture complex relationships between multiple financial attributes, especially when dealing with large and heterogeneous datasets. Modern financial datasets contain a wide range of customer information, including demographic, financial, and behavioural attributes. Conventional models may not effectively process such high-dimensional data, which can reduce prediction accuracy and limit the reliability of credit risk evaluation systems [6].

Furthermore, many existing loan evaluation systems rely on a limited number of data features that may not fully represent the financial behaviour and repayment capacity of loan applicants. As a result, these systems may incorrectly classify high-risk applicants as low-risk borrowers or reject applicants who are capable of repaying loans. Recent research has highlighted the importance of using advanced machine learning techniques to improve credit risk prediction accuracy and enable more efficient loan approval systems [2], [4].

Disadvantages Of The Existing System

- **Time-Consuming Decision Process**

Manual evaluation of loan applications requires significant time and effort, particularly when processing a large number of applicants in financial institutions.

- **Human Bias and Inconsistency**

Human decision-making processes may introduce biases and inconsistencies that affect the fairness and reliability of loan approval decisions [13].

- **Limited Data Processing Capability**

Traditional loan evaluation systems may not effectively analyse large financial datasets containing multiple customer attributes.

- **Lower Prediction Accuracy**

Conventional statistical models often fail to capture complex relationships between financial variables, resulting in less accurate credit risk predictions [1].

- **Inefficient Risk Assessment**

Existing systems may incorrectly classify loan applicants due to limited predictive capabilities and insufficient use of advanced data analytics techniques [2].

Proposed System

To overcome the limitations of traditional loan evaluation systems, this study proposes a machine learning-based loan risk prediction system that automatically analyses customer financial data and predicts the likelihood of loan default. The proposed framework utilizes historical loan application datasets containing multiple financial attributes to determine whether a loan applicant belongs to a low-risk or high-risk category. Machine learning techniques have demonstrated strong potential in improving credit risk prediction by identifying hidden patterns within large financial datasets [1], [2].

The proposed system uses historical loan application data that includes several customer attributes such as credit amount, checking account status, age, employment status, loan duration, and credit history. These attributes provide valuable insights into the financial behaviour and repayment capability of loan applicants. Financial institutions commonly analyse such attributes during credit evaluation processes to determine the creditworthiness of applicants and minimize the risk of loan defaults [3], [4].

The system begins with the data collection and preprocessing stage, which prepares the dataset for machine learning model training. Data preprocessing involves removing outliers, handling missing values, encoding categorical variables into numerical representations, and scaling feature values to maintain consistency across attributes. Additionally, the dataset is divided into training and testing subsets to ensure reliable evaluation of the predictive models. Proper preprocessing of financial datasets is essential to improve model accuracy and reduce bias during the training process [6].

After preprocessing, several machine learning algorithms are implemented to predict the credit risk associated with each loan application. The models considered in this study include Decision Tree, Random Forest, Support Vector Machine (SVM), Multi-Layer Perceptron (MLP) Neural Network, Naïve Bayes, and a Stacking Ensemble model. These algorithms analyse the financial attributes of loan applicants and classify them into low-risk or high-risk borrower categories. Previous studies have shown that machine learning classifiers such as Random Forest and ensemble models can significantly improve prediction accuracy in credit risk analysis by capturing complex relationships among financial variables [2], [8], [10].

To evaluate the performance of the implemented models, several performance metrics are used, including accuracy, precision, recall, F1-score, and confusion matrix analysis. These evaluation metrics provide a comprehensive assessment of model performance and help identify the most effective algorithm for credit risk prediction. Comparative analysis of multiple machine learning models allows financial institutions to select the most reliable predictive model for loan approval decision-making [7].

International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 2, Mar-Apr-2026, ISSN (Online): 2395-566X

By integrating machine learning techniques into the loan approval process, the proposed system provides several advantages over traditional loan evaluation methods. The automated framework enables faster processing of loan applications, reduces human bias in decision-making, and improves the overall prediction accuracy of credit risk assessment systems. Furthermore, the system can assist financial institutions in identifying the most influential financial factors that affect loan approval decisions, thereby improving risk management strategies and supporting more efficient lending practices [8], [13].

IV.SYSTEM DESIGN

System Architecture

Below diagram depicts the whole system architecture.

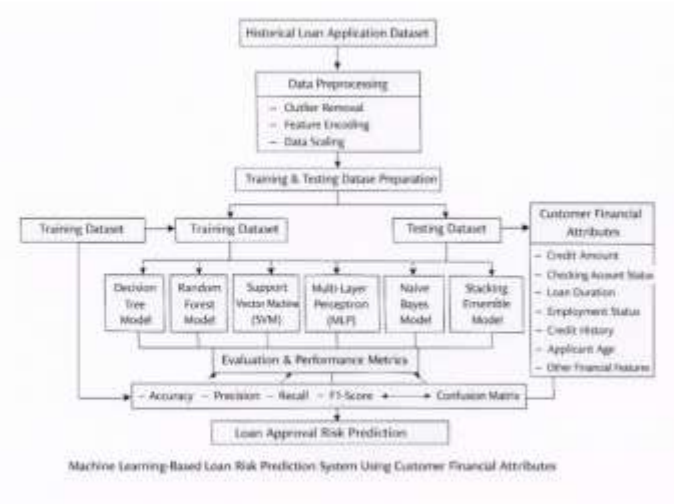

Fig 1. Methodology followed for proposed model

V. SYSTEM IMPLEMENTATION

Modules

Data Collection and Preprocessing

The first stage of the proposed system involves collecting a dataset containing customer financial and personal attributes used to evaluate loan applications. The dataset utilized in this study includes features such as checking account status, credit history, loan amount, loan duration, employment status, age, and other financial attributes. These features provide essential information for assessing the creditworthiness of loan applicants and predicting the probability of loan repayment. Publicly available financial datasets, such as credit risk datasets, are commonly used in machine learning research to develop and evaluate loan prediction systems [12].

Data preprocessing is performed to enhance the quality and reliability of the dataset before training machine learning models. Financial datasets often contain missing values, inconsistent records, and categorical attributes that must be converted into numerical form for model training. Therefore, preprocessing steps include identifying and removing outliers, encoding categorical variables into numerical values, and scaling feature values using normalization techniques. The dataset is subsequently divided into training and testing subsets to ensure reliable model evaluation and prevent overfitting during the training process [6].

Feature Selection and Data Analysis

In this module, the dataset is analysed to identify the most relevant features that influence loan risk prediction. Feature selection plays an important role in improving model performance by reducing data dimensionality and eliminating irrelevant attributes. Each feature in the dataset is examined to determine its contribution to loan approval decisions.

Data visualization techniques such as correlation heatmaps, bar charts, and distribution analysis are applied to explore relationships between variables and identify significant patterns within the dataset. These visualization methods help reveal correlations between financial attributes and loan repayment behaviour. Understanding these relationships allows the machine learning models to better classify loan applicants into low-risk or high-risk categories [9].

Training Machine Learning Models

Several machine learning algorithms are implemented to predict the risk level associated with loan applications. The models used in this study include Decision Tree, Random Forest, Support Vector Machine (SVM), Multi-Layer Perceptron (MLP) Neural Network, Naïve Bayes, and a Stacking Ensemble model. These algorithms have been widely applied in credit risk prediction systems due to their ability to analyse financial datasets and classify loan applicants based on their credit profiles [1], [3].

Each model is trained using the pre-processed training dataset. During the training phase, the algorithms learn patterns from historical loan data that help distinguish between applicants who are likely to repay their loans and those who may default. Additionally, hyperparameter tuning techniques are applied to optimize model performance and improve classification accuracy. Ensemble learning techniques, such as stacking models, combine multiple classifiers to enhance prediction performance and reduce classification errors [8], [10].

Loan Risk Prediction

Once the machine learning models are trained, they are used to evaluate new loan applications. The trained models analyse the applicant’s financial attributes and classify the loan request into either a low-risk or high-risk category. This automated prediction system assists financial institutions in evaluating loan applications quickly and efficiently.

International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 2, Mar-Apr-2026, ISSN (Online): 2395-566X

By reducing the reliance on manual evaluation processes, the system improves decision-making efficiency and minimizes human errors in credit risk assessment. Machine learning-based credit evaluation systems have been shown to significantly improve loan approval accuracy and support data-driven financial decision-making [2], [4].

Model Evaluation and Monitoring
The performance of each machine learning model is evaluated using several standard evaluation metrics, including accuracy, precision, and recall, F1-score, and confusion matrix analysis. These evaluation metrics provide detailed insights into the effectiveness of each model in correctly classifying loan applications and identifying potential defaulters.

Continuous monitoring of the prediction system allows financial institutions to periodically update and retrain the models using new financial data. This ensures that the prediction system remains accurate and adapts to evolving financial patterns and customer behaviour over time. Maintaining updated predictive models is essential for ensuring reliable credit risk assessment in dynamic financial environments [7].

VI. RESULTS AND DISCUSSION

To evaluate the performance of the proposed loan risk prediction system, multiple machine learning models were trained and tested using the prepared dataset containing customer financial attributes. The dataset was divided into training and testing subsets to ensure reliable model evaluation and to prevent overfitting during model training. Each model was evaluated using several performance metrics including accuracy, precision, recall, and F1-score, which are widely used in credit risk prediction research to measure classification performance [1], [3].

During the experimentation process, several machine learning algorithms were implemented, including Decision Tree, Random Forest, Support Vector Machine (SVM), Naïve Bayes, Multi-Layer Perceptron (MLP) Neural Network, and a Stacking Ensemble model. These algorithms analyse financial attributes such as checking account status, credit history, loan amount, loan duration, employment status, and age to determine whether an applicant belongs to a low-risk or high-risk category. Machine learning models have demonstrated strong potential in financial decision-making systems because they can identify hidden patterns within customer financial datasets [2], [4].

From Table 1, it can be observed that the Random Forest model demonstrated strong performance among the individual machine learning algorithms, achieving high accuracy and F1-score. Random Forest models are particularly effective because they combine multiple decision trees, which helps reduce overfitting and capture complex patterns within financial datasets [2].

Table 1: Performance Comparison of Machine Learning Models for Loan Risk Prediction

Model	Accuracy (%)	Precision	Recall	F1-Score
Decision Tree	86.4	0.84	0.83	0.83
Random Forest	91.2	0.90	0.89	0.89
Support Vector Machine	88.7	0.87	0.90	0.88
Naïve Bayes	85.6	0.84	0.89	0.86
MLP Neural Network	90.1	0.89	0.88	0.88
Stacking Ensemble (Proposed)	93.5	0.92	0.91	0.91

The Support Vector Machine (SVM) and Naïve Bayes models achieved strong recall values, indicating their capability to correctly identify a large number of low-risk loan applications. However, these models occasionally misclassified high-risk applicants as low-risk due to the imbalance present in the dataset. Class imbalance is a common challenge in credit risk prediction tasks and can affect the classification performance of certain algorithms [10].

The Stacking Ensemble model, which integrates predictions from multiple machine learning algorithms, achieved the highest overall accuracy of 93.5% among the evaluated models. Ensemble learning techniques improve predictive performance by combining the strengths of multiple classifiers and reducing prediction errors. Recent studies have shown that ensemble-based approaches can significantly enhance the reliability of credit risk prediction systems [8], [10].

ROC Curve Analysis
To further evaluate the classification performance of the proposed loan prediction models, a Receiver Operating Characteristic (ROC) curve analysis was conducted. The ROC curve illustrates the relationship between the True Positive Rate (TPR) and False Positive Rate (FPR) at different classification thresholds.

The ROC analysis indicates that the proposed Stacking Ensemble model achieved an Area Under the Curve (AUC) value of approximately 0.94, which demonstrates strong classification capability in distinguishing between low-risk and high-risk loan applicants. A ROC curve closer to the top-left corner of the graph represents a highly effective predictive model with a high true positive rate and a low false positive rate.

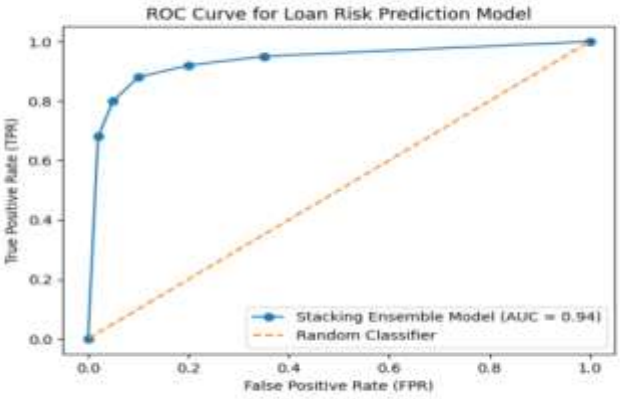

Fig. 2. ROC Curve for Loan Risk Prediction Model

In addition to performance metrics, feature importance analysis revealed that financial attributes such as credit amount, checking account status, age of the applicant, loan duration, and loan purpose play a significant role in determining the risk level of a loan application. These findings align with established credit evaluation practices used in the banking industry, where multiple financial indicators are analysed to assess borrower reliability [11].

Overall, the experimental results demonstrate that machine learning techniques can effectively support loan risk prediction and assist financial institutions in making data-driven lending decisions. By utilizing advanced machine learning and ensemble methods, financial institutions can improve loan approval accuracy, reduce the risk of loan defaults, and enhance the efficiency of credit risk management systems [8], [13].

VII. CONCLUSION AND FUTURE WORK

This study presented a machine learning-based loan risk prediction system designed to analyse customer financial attributes and determine the likelihood of loan default. The proposed framework utilizes several machine learning algorithms to classify loan applicants into low-risk and high-risk categories, thereby supporting financial institutions in making informed loan approval decisions. Machine learning techniques have demonstrated strong potential in financial risk assessment because they can identify hidden patterns within large financial datasets and improve the accuracy of credit risk prediction systems [1], [2].

The experimental results indicate that machine learning models significantly improve the efficiency and reliability of credit risk evaluation when compared to traditional manual assessment methods. Among the models evaluated in this study, the Random Forest and Stacking Ensemble models demonstrated superior performance, achieving higher accuracy and better classification capability for identifying high-risk borrowers. Ensemble learning techniques are particularly effective in financial prediction tasks because they combine the strengths of multiple algorithms to improve model robustness and reduce prediction errors [8], [10].

The proposed system can assist banks and financial institutions in reducing the risk of loan defaults by accurately identifying high-risk applicants before approving loans. By automating the loan evaluation process, the system also improves operational efficiency and enables data-driven decision-making in modern financial institutions. Machine learning-based loan prediction systems can enhance the overall credit evaluation process by providing faster and more consistent assessments of loan applications [3], [4].

For future work, the proposed system can be enhanced by incorporating larger and more diverse financial datasets to improve the generalization capability of the predictive models. Additionally, advanced machine learning approaches such as deep learning models, gradient boosting algorithms, and hybrid ensemble methods can be explored to further improve prediction accuracy and model robustness [7].

Integrating explainable artificial intelligence (XAI) techniques may also help improve transparency in automated loan approval systems by enabling financial institutions to better understand the factors influencing credit risk predictions [13]. Furthermore, future research should focus on addressing bias in training datasets and ensuring fairness in automated loan evaluation systems.

Developing ethical and transparent AI-based credit risk assessment models is essential to ensure that automated decision systems support fair lending practices while maintaining high predictive performance. By continuing to advance intelligent credit risk prediction systems, financial institutions can improve risk management strategies and enhance the overall reliability of loan approval processes.

REFERENCES

1. A. Archana, “A comparison of various machine learning algorithms and deep learning algorithms for prediction of loan eligibility,” International Journal for Research in Applied Science and Engineering Technology, vol. 11, no. 6, pp. 4558–4564, Jun. 2023, doi:10.22214/ijraset.2023.54495.
2. D. Dansana, S. G. K. Patro, B. K. Mishra, V. K. Prasad, A. R. Kaladgi, and A. W. Wodajo, “Analyzing the impact of loan features on bank loan prediction using Random

Forest algorithm," Engineering Reports, Jun. 2023, doi:10.1002/eng2.12707.
3. M. A. Sheikh, A. Goel, and T. G. Kumar, "An approach for prediction of loan approval using machine learning algorithm," in Proc. International Conference on Electronics and Sustainable Communication Systems (ICESC), Jul. 2020, doi:10.1109/ICESC48915.2020.9155614.
4. U. Orji, C. Ugwuishiwu, J. C. N. Nguemaleu, and P. N. Ugwuanyi, "Machine learning models for predicting bank loan eligibility," in Proc. IEEE Nigeria 4th International Conference on Disruptive Technologies for Sustainable Development (NIGERCON), Apr. 2022, doi:10.1109/NIGERCON54645.2022.9803172.
5. B. P. Lohani, M. Trivedi, R. J. Singh, V. Bibhu, S. Ranjan, and P. K. Kushwaha, "Machine learning based model for prediction of loan approval," in Proc. 3rd International Conference on Intelligent Engineering and Management (ICIEM), Apr. 2022, doi:10.1109/ICIEM54221.2022.9853160.
6. R. Karthiban, M. Ambika, and K. E. Kannammal, "A review on machine learning classification technique for bank loan approval," in Proc. International Conference on Computer Communication and Informatics (ICCCI), Jan. 2019, doi:10.1109/ICCCI.2019.8822014.
7. I. Awad, R. L. Ridwan, M. M. Muhammed, R. O. Abdulaziz, and G. A. Saheed, "Comparison of the CatBoost classifier with other machine learning methods," International Journal of Advanced Computer Science and Applications, vol. 11, no. 11, 2020, doi:10.14569/IJACSA.2020.0111190.
8. N. Uddin, M. K. U. Ahamed, M. A. Uddin, M. M. Islam, M. A. Talukder, and S. Aryal, "An ensemble machine learning based bank loan approval prediction system with a smart application," International Journal of Cognitive Computing in Engineering, vol. 4, pp. 327–339, Jun. 2023, doi:10.1016/j.ijcce.2023.09.001.
9. A. Shinde, Y. Patil, I. Kotian, A. Shinde, and R. Gulwani, "Loan prediction system using machine learning," ITM Web of Conferences, vol. 44, p. 03019, Jan. 2022, doi:10.1051/itmconf/20224403019.
10. S. Kokate and M. S. R. Chetty, "Credit risk assessment of loan defaulters in commercial banks using voting classifier ensemble learner machine learning model," International Journal of Safety and Security Engineering, vol. 11, no. 5, pp. 565–572, Oct. 2021, doi:10.18280/IJSSE.110508.
11. Wells Fargo, "Five Cs of Credit – What lenders look for." [Online]. Available: https:/ /www.wellsfargo.com/financial-education/credit-management/five-c/
12. "Credit Risk Customers Dataset," Kaggle, Apr. 12, 2023. [Online]. Available: https: //www.kaggle.com/datasets/ppb00x/credit-risk-customers
13. S. Townson, "AI can make bank loans more fair," Harvard Business Review, Nov. 6, 2020. [Online]. Available: https:/ hbr.org/2020/11/ai-can-make-bank-loans-more-fair