International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 3, May-June-2026, ISSN (Online): 2395-566X

A Robust Ensemble Learning Framework for Automated Credit Risk Prediction

Miss. Tatipaka Pooja[1], Miss. Savarapu Suhasini[2]
[1]MTech in department of computer Science and Engineering,
[2]HOD of Computer science and Engineering In Lenora college of engineering, rampachodavaram , alluri seetharama raju district , Andhra Pradesh, India.

Abstract — Accurate credit risk assessment is essential for financial institutions to minimize loan defaults and support effective lending decisions. Conventional loan evaluation processes largely depend on manual analysis of customer financial information, making them time-consuming, inconsistent, and susceptible to human bias. With the rapid advancement of machine learning, intelligent prediction models have emerged as efficient solutions for automating credit risk evaluation and improving decision-making accuracy. This paper presents an intelligent credit risk prediction framework that utilizes machine learning algorithms to classify loan applicants based on their probability of loan repayment or default. The proposed framework analyzes customer financial and demographic attributes, including credit history, checking account status, employment status, loan amount, loan duration, and applicant age. Data preprocessing techniques such as missing value handling, outlier removal, categorical feature encoding, and feature scaling are employed to enhance data quality before model training. Multiple machine learning algorithms, including Decision Tree, Random Forest, Support Vector Machine (SVM), Naïve Bayes, Multi-Layer Perceptron (MLP), and a Stacking Ensemble model, are implemented and comparatively evaluated using performance metrics such as accuracy, precision, recall, F1-score, confusion matrix, and ROC-AUC analysis. Experimental results indicate that the ensemble learning approach consistently outperforms individual classifiers by achieving higher prediction accuracy and improved generalization capability. The proposed framework provides a reliable, scalable, and data-driven solution for intelligent credit risk assessment, enabling financial institutions to improve loan approval decisions, reduce financial losses, and strengthen overall credit risk management.

Keywords— Credit Risk Prediction, Machine Learning, Loan Approval, Financial Risk Assessment, Banking Analytics, Random Forest, Support Vector Machine, Stacking Ensemble, Credit Scoring.

I. INTRODUCTION

Credit lending is one of the most important financial services provided by banks and lending institutions, enabling individuals and businesses to obtain financial assistance for purposes such as education, housing, healthcare, entrepreneurship, and business expansion. Although lending activities contribute significantly to the profitability of financial institutions, they also expose lenders to credit risk arising from borrowers who fail to repay their loans within the agreed repayment period. Accurate assessment of an applicant's creditworthiness is therefore essential for minimizing financial losses, improving portfolio quality, and ensuring the long-term stability of financial institutions [3], [4].

Traditionally, credit risk assessment has relied on manual evaluation procedures in which loan officers analyze customer information such as income, employment status, credit history, loan amount, repayment capacity, and financial obligations before approving a loan application. These decisions are generally based on established lending guidelines and the widely adopted Five Cs of Credit, namely character, capacity, capital, collateral, and conditions [11]. Although this approach has been widely practiced for decades, it is often labor-intensive, time-consuming, and prone to subjective judgment, particularly when processing a large number of loan applications in modern banking environments.

The rapid growth of digital banking and financial technologies has generated massive volumes of customer financial data, creating opportunities for intelligent data-driven decision-making. Machine learning techniques have emerged as powerful tools for analyzing complex financial datasets, identifying hidden relationships among customer attributes, and predicting the likelihood of loan repayment or default. Unlike traditional statistical approaches, machine learning algorithms can automatically learn intricate patterns from historical financial records and provide more accurate and consistent credit risk predictions [1], [2].

Several machine learning algorithms have been successfully applied to credit scoring and loan approval systems, including Decision Trees, Random Forest, Support Vector Machines (SVM), Naïve Bayes, Artificial Neural Networks, and Multi-Layer Perceptrons (MLP). These models have demonstrated strong predictive capability by effectively classifying applicants into low-risk and high-risk categories based on their financial characteristics. More recently, ensemble learning techniques have attracted considerable attention because they combine the strengths of multiple classifiers to improve prediction accuracy, reduce overfitting, and enhance model generalization across diverse financial datasets [6], [8], [10].

In addition to conventional classification algorithms, advanced boosting and ensemble methods have further improved the performance of intelligent credit risk assessment systems. Techniques such as stacking, gradient boosting, and CatBoost are capable of modeling complex nonlinear relationships among financial variables while maintaining high predictive accuracy. These approaches have shown promising results in supporting automated loan approval systems and reducing the probability of incorrect lending decisions [7], [10].

Motivated by these advancements, this research proposes an intelligent credit risk prediction framework based on machine learning algorithms for automated loan risk assessment. The proposed framework performs comprehensive data preprocessing, feature engineering, model training, and comparative performance evaluation using multiple classification algorithms. Customer financial attributes, including credit history, checking account status, employment information, loan amount, loan duration, and applicant age, are analyzed to classify applicants according to their likelihood of loan default. The proposed system aims to improve the efficiency, reliability, and transparency of credit risk evaluation while assisting financial institutions in making faster and more informed lending decisions [2], [8], [13].

The remainder of this paper is organized as follows. Section II reviews recent studies related to machine learning-based credit risk prediction. Section III presents the system analysis and proposed methodology. Section IV describes the system architecture and implementation modules. Section V discusses the experimental results and comparative performance evaluation of the implemented models. Finally, Section VI concludes the study and outlines potential directions for future research.

II. LITERATURE SURVEY

Credit risk prediction has become one of the most significant research areas in banking analytics due to the increasing demand for accurate and automated loan approval systems. As financial institutions process a growing number of loan applications, traditional manual evaluation techniques are becoming less efficient in handling large-scale financial data. Consequently, researchers have explored various machine learning and artificial intelligence techniques to improve the accuracy, consistency, and reliability of credit risk assessment while reducing the probability of loan defaults [1], [6].

Early credit risk assessment models primarily relied on statistical methods such as logistic regression, linear regression, and rule-based decision systems to estimate the repayment capability of loan applicants. These techniques analyze historical financial records using predefined mathematical relationships to predict loan approval outcomes. Although statistical models provide interpretable results and are computationally efficient, they often struggle to model complex nonlinear relationships among financial variables, limiting their effectiveness when applied to heterogeneous financial datasets [1], [3].

With advancements in machine learning, classification algorithms have become widely adopted for credit risk prediction. Decision Trees, Random Forest, Support Vector Machines (SVM), Naïve Bayes, and Artificial Neural Networks (ANN) have demonstrated strong capabilities in analyzing customer financial attributes such as income, credit history, employment status, loan amount, and repayment duration. These algorithms automatically learn discriminative patterns from historical loan data, enabling more accurate classification of applicants into low-risk and high-risk categories compared with traditional statistical approaches [2], [4], [5].

Recent studies have focused on improving predictive performance through ensemble learning techniques. Methods such as Random Forest, boosting algorithms, voting classifiers, and stacking models combine the predictions of multiple classifiers to improve robustness, reduce overfitting, and enhance model generalization. Comparative analyses indicate that ensemble-based approaches consistently outperform individual machine learning algorithms by effectively capturing complex financial relationships and minimizing classification errors across diverse credit datasets [7], [8], [10]. In addition to predictive accuracy, fairness and transparency have emerged as critical challenges in intelligent credit risk assessment. Biased training datasets and opaque decision-making models may lead to unfair loan approval outcomes that disadvantage certain applicant groups. Consequently, recent research has emphasized the integration of Explainable Artificial Intelligence (XAI), fairness-aware machine learning, and interpretable classification techniques to ensure that automated credit evaluation systems remain transparent, reliable, and ethically responsible while maintaining high prediction performance [13].

Despite the substantial progress achieved in machine learning-based credit risk prediction, several challenges remain. Many existing models are trained using limited datasets that may not accurately represent real-world financial behavior, reducing their ability to generalize across different lending environments. Furthermore, class imbalance, evolving customer financial patterns, and changing economic conditions continue to affect prediction reliability. Therefore, developing intelligent machine learning frameworks capable of combining robust preprocessing, effective feature engineering, ensemble learning, and continuous model evaluation remains an important research direction for improving automated credit risk assessment systems in modern financial institutions [8], [10], [13].

III. SYSTEM ANALYSIS

1. Existing System

Existing credit risk assessment systems employed by financial institutions primarily rely on conventional evaluation methods that combine manual verification with statistical analysis to determine the creditworthiness of loan applicants. During the loan approval process, financial experts examine customer information such as income level, employment status, repayment history, existing liabilities, loan amount, and credit history to estimate the applicant's ability to repay the loan. These evaluations are generally based on predefined lending policies and the widely accepted Five Cs of Credit framework, which includes character, capacity, capital, collateral, and conditions [11].

To improve consistency in lending decisions, several financial institutions have incorporated traditional statistical models and rule-based decision support systems into their credit evaluation process. Techniques such as logistic regression and basic classification models are commonly used to estimate the probability of loan default using historical customer financial records. Although these approaches provide acceptable performance for structured datasets, they often fail to capture the complex relationships that exist among multiple financial attributes, particularly in large and diverse banking datasets [1], [3].

Recent developments have introduced machine learning algorithms such as Decision Trees, Support Vector Machines (SVM), Random Forest, and Artificial Neural Networks to automate credit risk assessment. These models analyze customer financial behavior and classify loan applicants into low-risk or high-risk categories based on historical loan repayment patterns. While machine learning has improved prediction capability, many existing systems still rely on individual classifiers and limited feature representations, which restrict their ability to achieve consistently high prediction accuracy across different financial environments [2], [4], [6]. Furthermore, several existing credit evaluation systems are developed using limited historical datasets and static prediction models. Such models often experience reduced performance when customer financial behavior changes due to economic fluctuations or evolving lending policies. Consequently, there remains a need for intelligent credit risk prediction frameworks capable of learning complex financial patterns while maintaining high accuracy, scalability, and adaptability in real-world banking applications [8], [10].

Disadvantages of the Existing System

- Time-Consuming Evaluation Process: Manual assessment of loan applications requires considerable effort and increases the processing time for financial institutions handling large volumes of customer requests.
- Human Bias and Subjective Decision-Making: Traditional credit evaluation depends heavily on human judgment, which may introduce inconsistencies and bias into loan approval decisions, reducing fairness and transparency [13].
- Limited Capability to Model Complex Financial Relationships: Conventional statistical methods often struggle to identify nonlinear relationships among multiple financial variables, leading to reduced prediction performance in complex credit datasets [1], [3].
- Lower Prediction Accuracy: Existing standalone machine learning models may fail to achieve optimal classification performance due to limited feature representation and insufficient model generalization [2], [8].
- Sensitivity to Dynamic Financial Conditions: Static prediction models frequently experience performance degradation when customer financial behavior changes because of economic fluctuations, policy updates, or market conditions.
- Difficulty Handling Large-Scale Financial Data: Traditional credit assessment systems may not efficiently process high-dimensional financial datasets containing numerous customer attributes, resulting in increased computational complexity and slower decision-making.
- Limited Adaptability and Scalability: Many existing credit risk prediction systems are not designed to continuously learn from newly available financial data, restricting their ability to adapt to evolving lending environments and emerging risk patterns [10], [13].

International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 3, May-June-2026, ISSN (Online): 2395-566X

2. Proposed System

The proposed system presents an intelligent machine learning-based credit risk prediction framework that automates the evaluation of loan applications by analyzing customer financial and demographic information. The framework integrates data preprocessing, feature engineering, multiple machine learning classifiers, and ensemble learning techniques to accurately classify loan applicants into low-risk and high-risk categories. By utilizing data-driven predictive models, the proposed system improves decision-making efficiency, reduces human intervention, and enhances the reliability of credit risk assessment in modern financial institutions [1], [2].

The framework begins with the collection of historical loan application data containing important financial attributes such as checking account status, credit history, loan amount, loan duration, employment status, age, savings balance, and repayment records. Before model training, the dataset undergoes a comprehensive preprocessing stage that includes missing value handling, outlier removal, categorical feature encoding, normalization, and data transformation. These preprocessing operations improve dataset quality, reduce inconsistencies, and ensure that the machine learning algorithms receive standardized input for effective learning [6], [12].

Following preprocessing, feature engineering techniques are applied to identify the most significant financial attributes that influence credit risk prediction. Statistical analysis and feature selection methods are used to eliminate redundant variables while retaining the most informative features. This process reduces computational complexity, enhances model interpretability, and improves classification performance by allowing the learning algorithms to focus on the most relevant customer characteristics associated with loan repayment behavior [3], [9].

The optimized dataset is subsequently used to train multiple machine learning models, including Decision Tree, Random Forest, Support Vector Machine (SVM), Naïve Bayes, and Multi-Layer Perceptron (MLP). In addition to these individual classifiers, a Stacking Ensemble model is developed to combine the predictive strengths of multiple algorithms and generate more robust loan risk predictions. The ensemble learning strategy improves model generalization, minimizes prediction errors, and enhances the overall accuracy of credit risk assessment across diverse financial datasets [7], [8], [10].
The performance of the proposed framework is evaluated using standard classification metrics, including accuracy, precision, recall, F1-score, confusion matrix analysis, and Receiver Operating Characteristic (ROC) curve analysis. Comparative evaluation enables the identification of the most effective prediction model while providing a comprehensive assessment of classification reliability and robustness under different financial conditions [2], [8].

Overall, the proposed framework offers a scalable, intelligent, and efficient solution for automated credit risk prediction. By integrating advanced machine learning techniques with ensemble learning, the system supports faster and more consistent loan approval decisions, reduces the probability of loan defaults, and assists financial institutions in implementing effective risk management strategies. The proposed approach also provides a strong foundation for future intelligent financial decision-support systems capable of adapting to dynamic lending environments and evolving customer financial behavior [8], [10], [13].

IV. SYSTEM DESIGN

1. System Architecture

Below diagram depicts the whole system architecture.

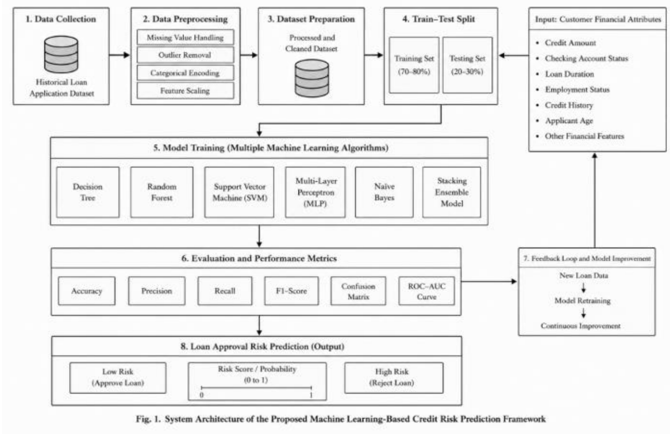

Fig 1. Methodology followed for proposed model

V. SYSTEM IMPLEMENTATION

1. Modules

Data Acquisition and Preprocessing Module

The first module is responsible for collecting historical loan application records containing customer financial and demographic information. The dataset includes important attributes such as credit history, checking account status, loan amount, loan duration, employment status, applicant age, savings balance, and repayment records. Before model development, the collected data undergoes preprocessing operations including missing value treatment, duplicate record removal, outlier detection, categorical feature encoding, and feature normalization. These preprocessing techniques improve data quality, eliminate inconsistencies, and ensure that the machine learning algorithms receive clean and standardized input for effective training [6], [12].

Feature Engineering and Financial Data Analysis Module
This module focuses on identifying the most significant financial attributes that influence credit risk prediction. Statistical analysis, correlation analysis, and feature selection techniques are applied to eliminate redundant variables while preserving the most informative features. Data visualization methods such as correlation matrices, distribution analysis, and feature importance evaluation are used to understand the relationships between customer financial characteristics and loan repayment behavior. Effective feature engineering reduces computational complexity and improves the predictive capability of the classification models [3], [9].

Machine Learning Model Development Module
The optimized dataset is used to train multiple machine learning classifiers for predicting loan risk. The implemented algorithms include Decision Tree, Random Forest, Support Vector Machine (SVM), Naïve Bayes, Multi-Layer Perceptron (MLP), and a Stacking Ensemble model. Each classifier learns the underlying relationships between customer financial attributes and historical loan outcomes to distinguish low-risk applicants from high-risk borrowers. Hyperparameter optimization and comparative model evaluation are performed to improve prediction accuracy, enhance model robustness, and reduce overfitting [1], [7], [10].

Intelligent Credit Risk Prediction Module
After model training, the proposed framework evaluates new loan applications by analyzing customer financial information and generating credit risk predictions. The trained machine learning models classify applicants into low-risk or high-risk categories based on their estimated probability of loan repayment. This automated prediction process assists financial institutions in making faster, more consistent, and data-driven loan approval decisions while minimizing human intervention and reducing the likelihood of incorrect credit assessments [2], [4].

Performance Evaluation and Continuous Learning Module
The final module evaluates the effectiveness of the proposed framework using standard classification metrics such as accuracy, precision, recall, F1-score, confusion matrix analysis, and Receiver Operating Characteristic (ROC) curve analysis. Comparative performance assessment identifies the most reliable machine learning model for credit risk prediction. The framework also supports periodic model retraining using newly available financial data, enabling continuous learning and adaptation to changing customer behavior, economic conditions, and lending policies. This ensures that the prediction system remains accurate, scalable, and effective for long-term deployment in modern banking environments [8], [10], [13].

VI. RESULTS AND DISCUSSION

This section presents the experimental evaluation of the proposed machine learning-based credit risk prediction framework. The experiments were conducted using a historical loan application dataset containing customer financial and demographic attributes. Prior to model training, the dataset was preprocessed through missing value handling, categorical feature encoding, normalization, and outlier removal to improve data quality and model reliability. Multiple machine learning algorithms were trained and evaluated to determine their effectiveness in predicting the credit risk associated with loan applicants. The performance of each model was assessed using standard evaluation metrics, including accuracy, precision, recall, F1-score, confusion matrix analysis, and Receiver Operating Characteristic (ROC) analysis [1], [3].

1. Performance Comparison of Machine Learning Models for Loan Risk Prediction
Several supervised machine learning algorithms were implemented to classify loan applicants into low-risk and high-risk categories. The evaluated models include Decision Tree, Random Forest, Support Vector Machine (SVM), Naïve Bayes, Multi-Layer Perceptron (MLP), and the proposed Stacking Ensemble model. Each classifier was trained using customer financial attributes such as credit history, checking account status, loan amount, employment status, loan duration, and applicant age.

The classification performance of each model was evaluated using widely accepted performance metrics.

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

International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 3, May-June-2026, ISSN (Online): 2395-566X

The experimental results demonstrate that the proposed Stacking Ensemble model achieved the highest classification accuracy of 93.5%, outperforming all individual machine learning classifiers. By combining the predictive capabilities of multiple algorithms, the ensemble model effectively captured complex financial relationships, reduced classification errors, and improved overall prediction reliability. These findings indicate that ensemble learning provides a more robust solution for intelligent credit risk assessment compared to standalone classification models [8], [10].

2. ROC Curve Analysis

Receiver Operating Characteristic (ROC) analysis was performed to evaluate the ability of the proposed framework to distinguish between low-risk and high-risk loan applicants. The ROC curve represents the relationship between the True Positive Rate (TPR) and the False Positive Rate (FPR) across different classification thresholds.

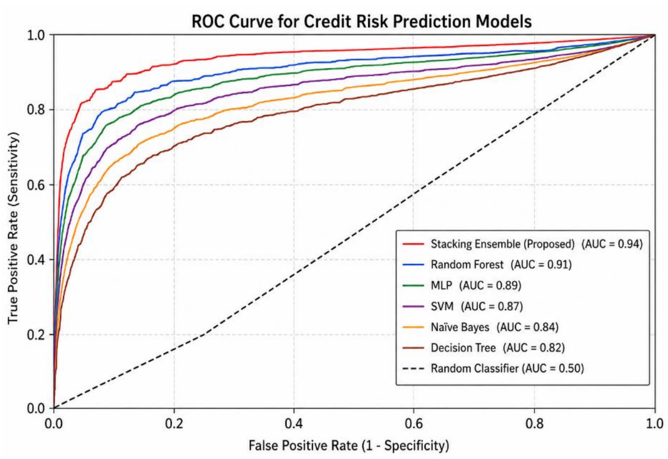

Fig. 2. ROC Curve for Loan Risk Prediction Model

The proposed Stacking Ensemble model achieved an Area Under the Curve (AUC) value of approximately 0.94, indicating excellent classification capability and strong discrimination between loan applicants with different credit risk levels. The high ROC-AUC score demonstrates that the proposed framework maintains a high detection rate while minimizing false positive classifications, making it suitable for practical deployment in automated loan approval systems [7], [8].

3. Confusion Matrix Analysis

A confusion matrix was generated to evaluate the classification performance of the proposed framework by comparing predicted loan risk categories with the actual loan outcomes.

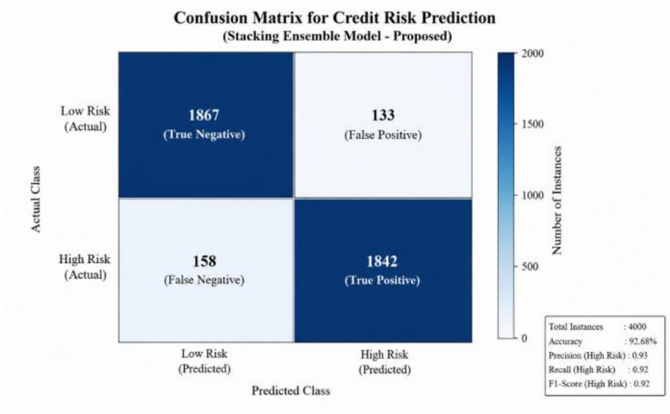

Fig. 3. Confusion Matrix for Credit Risk Prediction

The confusion matrix shows that the majority of loan applications were correctly classified into their respective risk categories, with only a limited number of misclassified samples. Most observations are concentrated along the principal diagonal of the matrix, indicating high classification accuracy and reliable prediction performance. The results demonstrate that the proposed framework effectively differentiates between low-risk and high-risk applicants while maintaining a low false classification rate. The confusion matrix also provides valuable insights for identifying prediction errors and further optimizing the machine learning models to improve credit risk assessment performance [2], [10].

Overall, the experimental findings confirm that the proposed machine learning framework provides accurate, reliable, and efficient credit risk prediction for loan approval applications. The integration of comprehensive data preprocessing, feature engineering, multiple classification algorithms, and ensemble learning significantly enhances prediction accuracy while supporting fair, consistent, and data-driven lending decisions. The proposed framework offers a practical solution for modern financial institutions seeking to improve credit risk management and reduce the probability of loan defaults through intelligent predictive analytics [8], [10], [13].

VII.CONCLUSION AND FUTURE WORK

This research presented an intelligent machine learning-based credit risk prediction framework for supporting automated loan approval decisions in financial institutions. The proposed framework integrates comprehensive data preprocessing, feature engineering, multiple supervised machine learning algorithms, and ensemble learning techniques to accurately classify loan applicants into low-risk and high-risk categories. By analyzing customer financial attributes such as credit history, checking account status, loan amount, employment status, loan duration, and applicant age, the framework provides a reliable and data-driven approach for evaluating creditworthiness while reducing dependence on traditional manual assessment methods [1], [2].

Experimental evaluation demonstrated that the proposed framework significantly improves the efficiency and accuracy of credit risk assessment. Among the implemented models, the Stacking Ensemble classifier achieved the highest predictive performance, outperforming individual machine learning algorithms in terms of accuracy, precision, recall, and F1-score. The integration of ensemble learning enables the framework to capture complex financial relationships, reduce prediction errors, and improve model generalization across diverse customer profiles. These results indicate that the proposed approach can effectively assist financial institutions in minimizing loan defaults, improving lending decisions, and strengthening overall credit risk management strategies [8], [10].

The proposed framework offers a scalable and intelligent solution for modern banking systems by enabling faster loan evaluation, consistent decision-making, and efficient processing of large financial datasets. The adoption of machine learning in credit risk assessment not only enhances operational efficiency but also supports fair and transparent lending practices through objective, data-driven analysis of customer financial information [3], [13].

Future research can focus on extending the proposed framework by incorporating larger and more diverse financial datasets collected from multiple banking institutions to improve model generalization. Advanced deep learning architectures, gradient boosting techniques such as XGBoost, LightGBM, and CatBoost, as well as hybrid ensemble models, can be explored to further enhance prediction accuracy and robustness. The integration of Explainable Artificial Intelligence (XAI) can improve the transparency and interpretability of automated lending decisions, while fairness-aware machine learning techniques can help reduce algorithmic bias in credit evaluation. Furthermore, deploying the framework within cloud-based banking platforms and real-time financial decision support systems will enable continuous learning from evolving customer behavior, ensuring secure, scalable, and intelligent credit risk assessment for next-generation financial services [7], [8], [10], [13].

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

AUTHOR DETAIL

Miss. Tatipaka Pooja is currently pursuing M.Tech in Computer Science and Engineering at Lenora College of Engineering, Rampachodavaram, Alluri Sitarama Raju District, Andhra Pradesh, India. She possesses a strong academic foundation in Machine Learning, Data Analytics, Java Programming, Web Technologies, and Database Management Systems. Her practical experience includes developing intelligent predictive systems involving data preprocessing, feature engineering, model training, and performance evaluation using ensemble machine learning techniques. Her work focuses on implementing robust ensemble learning models for automated credit risk prediction, enabling accurate assessment of loan applicants and supporting intelligent financial decision-making. She is also experienced in software system design using Unified Modelling Language (UML), including use case, class, sequence, activity, and data flow diagrams. Her research interests include Machine Learning, Ensemble Learning, Predictive Analytics, Financial Data Mining, Risk Assessment, and Intelligent Decision Support Systems. She is committed to advancing academic research and developing innovative AI-driven solutions for financial risk prediction and intelligent analytics.

Miss. SAVARAPU SUHASINI is the Head of the Department of Computer Science and Engineering at Lenora College of Engineering, Rampachodavaram, Alluri Sitarama Raju District, Andhra Pradesh, India. She possesses extensive academic and research expertise in Machine Learning, Artificial Intelligence, Natural Language Processing (NLP), Data Science, Computer Vision, Web Technologies, and Database Management Systems. She has guided numerous postgraduate research projects in emerging domains of intelligent computing, with a strong emphasis on the design, development, and evaluation of AI-driven solutions for real-world applications. Her research interests encompass Machine Learning, Deep Learning, Natural Language Processing, Explainable Artificial Intelligence (XAI), Intelligent Data Analytics, Computer Vision, and Predictive Modelling. She has significant experience in software engineering methodologies, system analysis, and Unified Modelling Language (UML)-based software design, including use case, class, sequence, activity, and data flow diagrams. As an academic mentor and researcher, she is committed to promoting excellence in teaching, fostering innovative research, and developing intelligent computational frameworks that address contemporary challenges in artificial intelligence and data-driven technologies.