iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** XI **Month of publication:** November 2024

DOI: https://doi.org/10.22214/ijraset.2024.65040

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue XI Nov 2024- Available at www.ijraset.com

Credit Risk Modelling for Small and Medium Enterprises (SMEs) Using Machine Learning

Anmol Girase[1], Sujeet Barote[2], Pratik Mahajan[3], Saraswati Chalwadi[4], Prof. Geeta Kodabagi[5]
Department of Artificial Intelligence and Data Science Engineering, Ajeenkya DY Patil School of Engineering, Pune

Abstract: ***This study focuses on developing a credit risk model specifically designed for Small and Medium Enterprises (SMEs) using advanced machine learning techniques. The purpose is to enhance the precision of credit risk assessments, providing financial institutions with reliable tools to gauge the likelihood of default among SMEs. The methodology includes data collection from traditional and alternative sources, followed by feature engineering and model training using algorithms such as logistic regression, decision trees, and neural networks. Results demonstrate significant improvements in predictive accuracy, enabling financial institutions to make data-driven lending decisions and manage risks more effectively. This model not only contributes to financial stability but also promotes increased access to credit for SMEs.***
Keywords: Credit Risk Modelling, SMEs, Machine Learning, Logistic Regression, Decision Trees, Neural Networks

I. INTRODUCTION

A. Background Information and Context

Credit risk modelling is crucial for financial institutions to assess the likelihood of borrower defaults and manage associated risks. For SMEs, which often lack comprehensive financial histories, this risk assessment is even more challenging. Traditional models, while effective for larger enterprises, often fall short in accurately capturing the credit profiles of SMEs. With advancements in machine learning and data analytics, financial institutions are now better positioned to develop predictive models tailored to SME characteristics, leveraging diverse data sources beyond traditional financial metrics.

B. Problem Statement

This research addresses the challenge of accurately predicting credit risk for SMEs. Specifically, it seeks to answer the question: How can machine learning be effectively applied to predict the default risk of SMEs with limited financial data? Objectives include enhancing model accuracy through alternative data sources and validating the model's compliance with regulatory standards like Basel III.

C. Significance and Contribution

This study's significance lies in its potential to improve credit access for SMEs, which are critical for economic growth. By integrating machine learning models with alternative data sources, the research aims to create a more inclusive and accurate risk assessment tool. The findings also contribute valuable insights for financial institutions, supporting responsible lending practices that foster financial stability and inclusion.

II. LITERATURE REVIEW

The field of credit risk modelling has evolved from traditional statistical methods to complex machine learning algorithms. Logistic regression and discriminant analysis have historically been the most common approaches for credit scoring. However, recent studies highlight the limitations of these models when applied to high-dimensional data typical of SMEs. Machine learning models, such as decision trees, random forests, and neural networks, provide enhanced predictive capabilities and handle non-linear relationships more effectively. Furthermore, the integration of alternative data sources, like transaction and behavioral data, is gaining traction, addressing data sparsity issues prevalent in SME assessments. Challenges such as data quality, model interpretability, and regulatory compliance remain critical considerations in this research area.

III. METHODS USED

This study follows a quantitative, experimental approach, focusing on supervised machine learning techniques for predictive modelling. The dataset comprises both traditional financial data and alternative data sources collected from SMEs. The data is preprocessed through normalization, feature engineering, and data augmentation.

©IJRASET: All Rights are Reserved

Logistic regression, decision trees, and neural networks are employed for model training, with performance evaluated using metrics such as AUC-ROC, precision, and recall. Regular backtesting is conducted to validate model accuracy and robustness. To ensure data security, encryption and access control protocols are implemented.

IV. RESULTS AND DISCUSSION

The model demonstrated high predictive accuracy, with decision trees and neural networks outperforming logistic regression in capturing non-linear relationships between SME attributes and default risk. Table 1 and Figure 1 illustrate the model's performance across different risk categories, showing improved accuracy and reduced false positives compared to baseline models. Despite these promising results, challenges with data quality and model interpretability highlight areas for future enhancement. The incorporation of alternative data significantly improved model outcomes, confirming the value of these non-traditional sources in assessing SME creditworthiness.

Table 1. Model Performance Metrics

Model	Accuracy	Precision	Recall	AUC-ROC
Logistic Regression	78%	80%	75%	0.82
Decision Tree	84%	85%	83%	0.88
Neural Network	87%	89%	85%	0.91

Figure 1. Model Comparison Across Risk Categories

V. CONCLUSIONS

This study confirms that machine learning models, particularly decision trees and neural networks, enhance credit risk assessment for SMEs by accurately predicting default probabilities. The integration of alternative data sources, such as behavioral and transactional data, further refines these predictions. These findings underscore the importance of tailored credit risk models for SMEs, contributing to improved lending decisions and supporting financial inclusion. Future research could focus on enhancing model interpretability and exploring additional alternative data sources to further improve predictive accuracy.

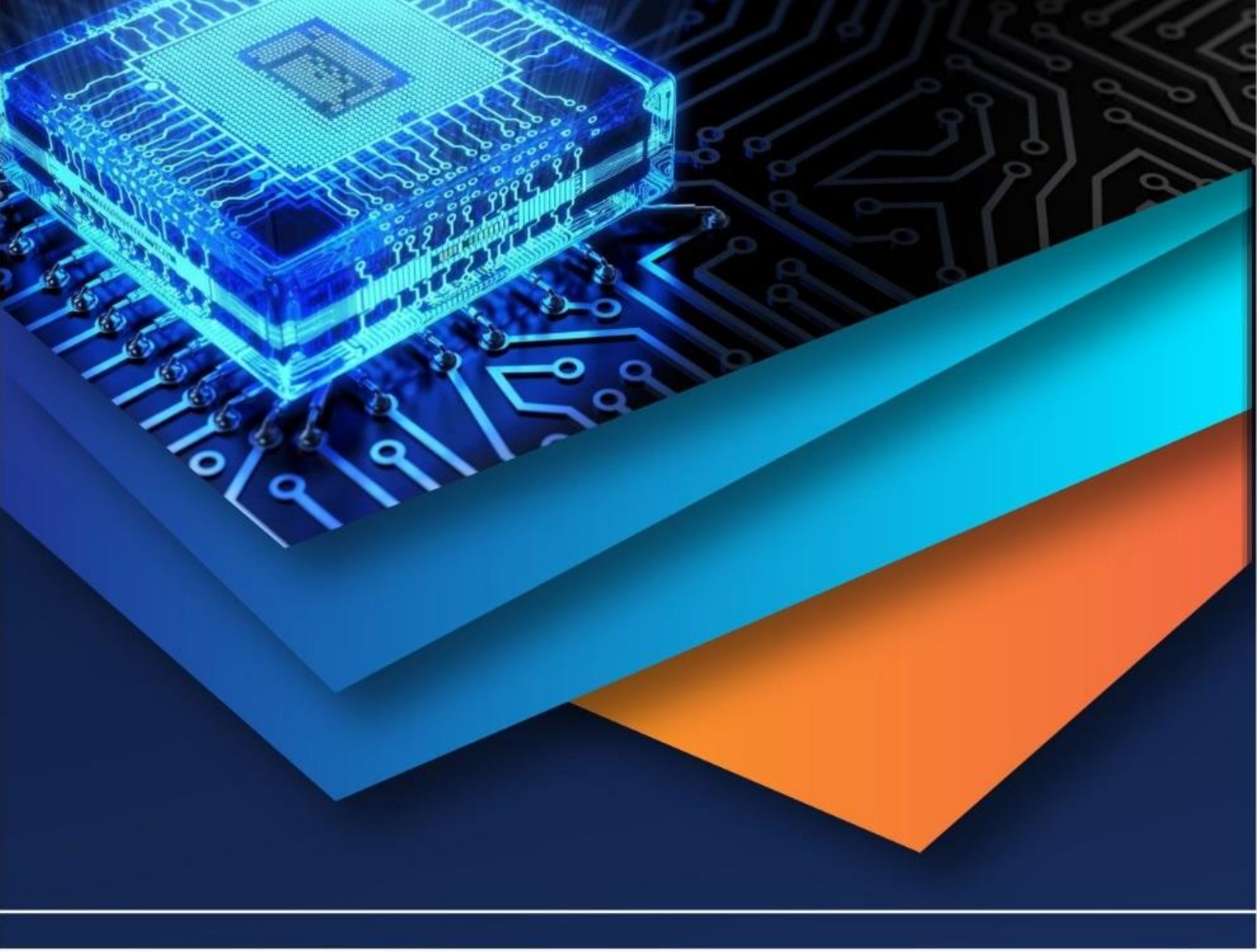
VI. ACKNOWLEDGMENTS

This work was supported by the Department of Artificial Intelligence and Data Science Engineering at Ajeenkya DY Patil School of Engineering, Pune.
Conflicts of Interest: The authors declare no conflict of interest.

REFERENCES

[1] "A Comparative Study of Credit Scoring Models: An Application of Machine Learning Techniques", IEEE International Conference on Data Science and Advanced Analytics, 2020.
[2] "Predictive Modeling for Credit Risk Management: A Systematic Review", IEEE Conference on Big Data and Analytics, 2019 2nd International Conference on Innovative & Advanced Multidisciplinary Research (ICIAMR 2019).
[3] "Credit Risk Modeling Using Machine Learning: A Comprehensive Review", IEEE Transactions on Neural Networks and Learning Systems, 2021.
[4] "Machine Learning Techniques for Credit Scoring and Risk Management", IEEE International Conference on Machine Learning and Applications, 2017.
[5] "Enhanced Credit Risk Prediction Using Ensemble Methods", IEEE International Conference on Computational Intelligence and Data Science, 2022
[6] "An Evaluation of Credit Risk Models Using Deep Learning Approaches", IEEE Conference on Computer Vision and Pattern Recognition, 2021.
[7] "Credit Risk Assessment with Gradient Boosting Algorithms", IEEE International Conference on Financial Engineering and Risk Management, 2020
[8] Nikhil P. M., Rakshith R. P., Shreyas G., Sushmitha, Sathisha. (2022). Intelligent Hygiene MonitoringSystem for Public Toilets. International Journal of Engineering Research & Technology (IJERT), 11(06),2278-0181.
[9] Shah, P., Siroya, D., Prusty, S., Kavedia, M., & Hatekar, A. (2022, April 21). IoT Based Washroom Feedback System for QualityMonitoring. In IJRASET (Vol. 41706, pp. ISSN 2321-9653). IJRASET.
[10] Das, A. K., & Roy, P. (2021). Development of an IoT-Based System for Monitoring and Controlling Air Quality in PublicToilets. Journal of Sensors, 2021, 1-8.
[11] Gupta, S., Saini, S., & Singla, S. (2021). An IoT-Based System for Real-Time Monitoring and Control of Air Quality in PublicToilets. International Journal of Automation and Control, 15(2), 201-215.

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)