Unleashing the Power of Machine Learning: A Comparative Study of Classification Algorithms for Credit Risk Assessment

Sandeep Shinde[1] and Satish Kale[2]

Lecturer, Bharati Vidyapeeth Institute of Technology, Navi Mumbai, India[1,2]

Abstract: *: Credit risk assessment is a critical task in the financial industry, enabling lenders to evaluate the likelihood of default and make informed lending decisions. This paper presents a comprehensive comparative study of classification algorithms for credit risk assessment using machine learning techniques. The paper commences by providing an overview of the importance of credit risk assessment and the challenges faced by traditional methods. It then delves into the exploration of various machine learning algorithms, including logistic regression, decision trees, random forests, support vector machines, and neural networks, highlighting their potential in credit risk assessment. The objective of this study is to compare the performance of different classification algorithms in credit risk assessment. To achieve this, a dataset comprising historical credit data, including borrower information, financial indicators, and repayment history, is collected. The dataset is preprocessed to handle missing values, outliers, and feature engineering is applied to extract relevant predictors. A comprehensive evaluation is conducted, considering performance metrics such as accuracy, precision, recall, and area under the receiver operating characteristic curve (AUC-ROC). The comparative study provides insights into the strengths and weaknesses of each algorithm and their suitability for credit risk assessment in different scenarios. The results of this study contribute to the existing literature on credit risk assessment and offer practical guidance for financial institutions in selecting appropriate machine learning algorithms. Furthermore, the paper discusses potential challenges and limitations associated with the application of machine learning in credit risk assessment and proposes future research directions.*

Keywords: Credit Risk Assessment, Machine Learning, Classification Algorithms

REFERENCES

[1]. Paleologo, G., Elisseeff, A. and Antonini, G., Subagging for Credit Scoring Models. European Journal of Operational Research, 201 (2010), 490-499

[2]. Hong, S. K. and Sohn, S. Y., Support vector machines for default prediction of SMEs based on technology credit. European Journal of Operational Research, 201 (2010), 838-846

[3]. Khandani, A. E., Kim, A. J. and Lo, A. W., Consumer credit-risk models via machine-learning algorithms. Journal of Banking & Finance, 34 (2010), 2767-2787

[4]. Khashman, A., A Neural Network Model for Credit Risk Evaluation. International Journal of Neural Systems, 19 (2009), 285—294

[5]. Van Sang Ha, Ha Nam Nguyen and DucNhan Nguyen, 2016. A novel credit scoring prediction model based on Feature Selection approach and parallel random forest, Indian Journal of Science Vol. 9(20).

[6]. NazeehGhatasheh, 2014. Business Analytics using Random Forest Trees for Credit Risk Prediction: A Comparison Study, International Journal of Advanced Science and Technology.