Chapter 18
Performance Evaluation of Different Machine Learning Algorithms Using Credit Scoring Model

Amrit Singh
NIST Institute of Science and Technology (Autonomous), India

Harisankar Mahapatra
NIST Institute of Science and Technology (Autonomous), India

Anil Kumar Biswal
https://orcid.org/0000-0001-7341-216X
Udayanath College of Science and Technology (Autonomous), India

Milan Samantaray
Trident Academy of Technology, India

Debabrata Singh
Institute of Technical Education and Research, Siksha 'O' Anusandhan (Deemed), India

ABSTRACT

The project focuses on the development of a credit scoring model. Concerns with credit scoring are being raised when developing an empirical model to support the financial decision-making process for financial institutions. This chapter focuses on the development of a credit scoring model using a combination of feature selection and ensemble classifiers. The most relevant features are identified, and an ensemble classifier is used to reduce the risk of overfitting with the aim of improving the classification performance of credit scoring models in the proposed method. Several metrics, including accuracy, precision, recall, F1 score, and AUC-ROC, are used to evaluate the performance of the model. The accuracy and robustness of credit scoring models can potentially be improved by the proposed method, and the evaluation metrics can be used to further enhance it.

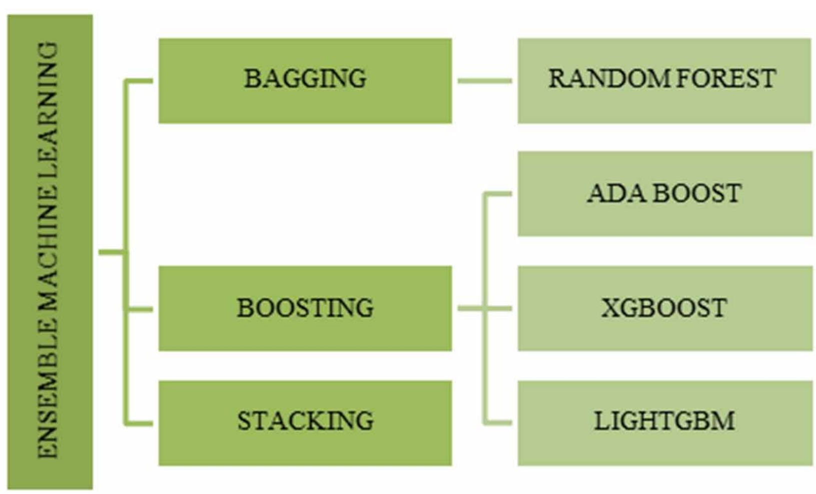
DOI: 10.4018/978-1-6684-9809-5.ch018

INTRODUCTION

In the banking system, credit scoring is a process used by lenders to evaluate the creditworthiness of potential borrowers. The likelihood of defaulting on a loan is determined by analyzing the borrower's credit history, financial information, and other relevant data as part of the process (Li, Y. et al. 2020). The importance of credit scoring for the industrial and banking systems cannot be overstated, as even a small improvement of 1% or 2% in accurately recognizing applicants with bad credit can result in significant savings for financial institutions (Gunnarsson, B. R. et al. 2021). Originally, credit scoring was evaluated subjectively based on personal experiences.

However, in today's world, with the explosion of data, classical statistical analysis models' elastic performance is not very good when it comes to handling large quantities of data. Consequently, the accuracy of the predictions is affected as some assumptions in these models cannot be established (Luo, C. et al. 2017) (Khalili, N. et al. 2023). With the advent of machine learning techniques and Ensemble learning, credit scoring has undergone a transformational change, enabling the development of more accurate and efficient credit risk models (Jiang, C. et al. 2023) (Xu, C. et al. 2023). Scoring calculations are based on a customer's payment records, frequency of payments, amount of debts, credit charges-offs, and other transaction activities (Asencios, R. et al. 2023) (Reji, T. et al. 2023).

Figure 1. Ensemble machine learning framework



MOTIVATION

Motivation of this credit score model is to help the financial institution find defaulter and easy evaluation of credit score. There are billions of dollar transaction going around the world. This underscores the economic value of credit scoring models, which are crucial in assessing a borrower's creditworthiness. Credit scoring in the banking system is essential for making informed lending decisions, reducing the risk of default, and ensuring the financial stability of the lending institution. By using credit scoring models, banks can determine the appropriate interest rate, credit limit, and repayment terms for each