ABSTRACT

Exploring the intersection of social media and financial technology, we propose a credit scoring model utilizing LinkedIn's social media data to enhance the assessment of individual creditworthiness through an ensemble soft voting classifier. The study aims to address the issue of creditworthiness for people without traditional credit histories. Using social media data increases access to credit and enhances performance through the ensemble soft voting technique. The ensemble approach combines multiple machine learning algorithms, including Decision Tree, Naive Bayes, Logistic Regression, Support Vector Machine, and Random Forest to enhance predictive performance. Our results from the proposed method demonstrate good performance. The predictions from each classifier model are combined by calculating a weighted average of the predicted probabilities for each class. Compared to the Decision Tree classifier that has similar performance, the ensemble soft voting method performs better in terms of ROC/AUC value indicating the effectiveness of the classifier for credit scoring. This research demonstrates the potential of using social media data to improve credit accessibility and showcase innovative applications of machine learning in finance.

Keywords: Credit Scoring Model, Voting Classifier, Machine Learning, and Social Media Data.