ABSTRACT

Flight delays pose a significant challenge in the aviation industry, affecting passenger satisfaction, airline operational efficiency, and airport performance. Although various managerial strategies and operational policies, such as airport capacity management systems and flight pattern analysis, have been implemented, they often fall short in delivering adaptive and precise solutions. This study focuses on binary classification of flight delays by employing the stacking ensemble learning method to improve predictive performance. Additionally, it incorporates hyperparameter optimization using Optuna to enhance model efficiency. The dataset, sourced from the United States Bureau of Transportation Statistics, consists of 1,800,000 rows and 120 features spanning from January to December 2024. Five models were tested: Decision Tree, Naïve Bayes, LightGBM, XGBoost, and a stacking ensemble of LightGBM and XGBoost. The results demonstrate that the stacking model outperformed the others, achieving a test accuracy of 0.8213 and an AUC score of 0.7564. Compared to the lowest-performing single model, stacking improved accuracy by 0.0021 and AUC by 0.0006. These findings highlight the effectiveness of stacking ensemble learning in significantly enhancing the classification performance of flight delay prediction models.

Keywords : Flight Delay, Machine Learning, Stacking Ensemble Learning, Optuna, Classification