ABSTRACT

Customer churn is the tendency of customers to leave a company or switch from one company to another company. High customer churn must be accurately classified to help companies predict customers who have the potential to churn. One approach to predicting customer churn is the supervised learning approach. Supervised learning consists of two problems, namely regression and classification. The supervised learning method used is gradient boosting, which handles classification problems. Gradient boosting combines the predictions of weak learners into strong learners by correcting previous predictions. The performance of the resulting model can be improved with GridSearchCV as a hyperparameter tuning. GridSearchCV selects the best combination of hyperparameters by testing each combination and validating each combination with a cross-validation technique. The study shows that gradient boosting with GridSearchCV for customer churn prediction at Bank ABC (Arab Banking Corporation), using stratified k-fold cross-validation with k = 5, results in an accuracy of 87% and an AUC of 0,87. This value explains that the model can effectively classify and predict class 1 (churn) and class 0 (no churn).

Keywords: Customer Churn, Gradient Boosting, GridSearchCV