ABSTRACT

Airplanes are one of the transportation that is in great demand by the public because of their superiority in travel time efficiency. It is undeniable that airline ticket prices often change due to several factors such as the travel season, seat availability, ticket purchase time, and so on. Monitoring and observation alone do not allow one to predict flight prices accurately, a prediction system is needed that is able to predict prices accurately. This study uses data on airline ticket prices on the Jakarta (CGK) – Denpasar (DPS) flight route taken from the Traveloka ticket booking website. Prediction is carried out using three regression algorithms, namely *XGBoost Regression, CatBoost Regression*, and *Random Forest Regression.* The results of this study show that the *CatBoost Regression* algorithm produces the highest evaluation value R^2 of 79.2% and the smallest MAE value of 0.202. *XGBoost Regression* before *tuning* resulted in an evaluation of 70.3% and an MAE of 0.237. Meanwhile, R^2 *XGBoost Regression* after tuning resulted in an evaluation of 75.6% and an MAE of 0.224. R^2 Random Forest Regression resulted in an evaluation of 74.7% and an MAE of 0.207. The analysis of the influence of variables using Shapley Additive Explanations (SHAP) which shows that the three highest variables that affect the model are the Durasi, Garuda Indonesia, and Transit variables. This study provides important insights into the factors that affect airline ticket prices and the potential use of regression algorithms to accurately predict prices.

Keywords *:* Airline Price Prediction, XGBoost Regression, CatBoost Regression, Random Forest Regression, Ensemble Method.