ABSTRACT

Online transportation services such as uber and lyft have become one of the basic human needs because it makes it easier for humans to carry out their activities. As time goes by, the data generated by uber and lyft ride users is increasing, such as consumer travel history data, prices for each trip taken by consumers, etc. With this data, researchers use it to predict the price of the trip before the start of the trip to be taken by consumers, so that consumers can estimate and prepare a budget in advance. In this study, researchers used two models, namely linear regression and random forest. Of the two modeling, the results of mean absolute error (MAE), mean square error (MSE), accuracy/mean absolute percentage error (MAPE), and explained variance score of random forest modeling are better than linear regression. In the random forest model, the accuracy/MAPE (mean absolute percentage error) result was obtained by 0.97, mean absolute error by 1, mean square error by 2.6, and explained variance score of 0.97.

Keywords: Linear regression, Random forest, Machine learning, Pricing Model