Airbnb Price Prediction Using Machine Learning

Quanfang Tan

The Affiliated International School of Shenzhen University, Shenzhen China

Abstract

Over the past few decades, the sharing economy has dramatically changed the way people live across many areas, including finance, travel, and hospitality. Airbnb is one of the largest rental platforms, playing a critical role in the sharing economy all over the world. This study examines the dynamic pricing on Airbnb by analyzing data on 2023 in New York City, US, with 39,626 Airbnb listings and each is described by 76 features. Four different machine learning methods -- linear regression, random forest (RF), extreme gradient boosting (XGBoost), and long short-term memory (LSTM) -- are utilized to identify key influencing features and predict the optimal price. The performance of those methods is measured based on two evaluation metrics -- R-squared (R^2) and root mean square error (RMSE). Our results show that RF outperforms other models in the training set given highest R^2 of 0.9490 and lowest RMSE of 0.1703, while XGBoost is the best model among these four models in the test set with the highest R^2 of 0.5456 and lowest RMSE of 0.5015. The findings indicate that ML is applicable in Airbnb listings pricing problems, and this also contributes insights for future pricing strategies on Airbnb, and offer valuable recommendations for stakeholders.

Keywords: Sharing Economy, Airbnb Pricing, Machine Learning