

ABSTRACT

The growth of the used smartphone market is driven by high consumer demand for affordable devices. Variations in smartphone specifications create challenges in determining appropriate pricing. This study applies the random forest regression algorithm to predict used smartphone prices based on device features. The dataset consists of 869 used smartphone records, divided into 608 training data and 261 testing data. Hyperparameter tuning was performed using GridSearchCV with repeated k-fold cross-validation (10 folds, 4 repetitions) to obtain the optimal parameter combination. The resulting random forest regression model achieved a MAPE value of 9.53%, indicating excellent predictive performance as it falls below 10%. This accuracy surpasses that of the default parameters, which yielded a MAPE of 10.28%. Variable importance analysis revealed that the most influential feature in price prediction is the network type (GSM/CDMA/HSPA/EVDO/LTE/5G) with an importance score of 50.46%. Further analysis using only the top five most important features did not guarantee better accuracy, as the MAPE remained around 9%, but it improved model efficiency. This indicates that using all 11 features provides similar performance to using only the top five most important ones. This study highlights that the random forest algorithm is effective for estimating used smartphone prices based on relevant device features, with proper parameter tuning ensuring optimal predictive performance.

Keywords: Random Forest, GridSearchCV, Price Prediction, Used Smartphones, Regression