

ABSTRAK

Kanaya Tabhita Djie (01112180013)

PREDIKSI HARGA TIKET PESAWAT MASKAPAI PENERBANGAN DI INDONESIA MENGGUNAKAN *TREE-BASED MODELS*

Skripsi, Fakultas Sains dan Teknologi (2023)

(xvi + 92 halaman, 47 gambar, 25 tabel, 2 lampiran)

Setiap harinya, data yang dihasilkan pada industri penerbangan Indonesia sangat besar. Data yang tersedia seperti jam keberangkatan, jam kedatangan, jumlah transit, hari penerbangan, dan sebagainya dapat digunakan untuk memprediksi harga tiket pesawat. Penelitian ini menggunakan tiga metode *machine learning* yaitu *Regression Tree*, *Bagging Regression Tree*, dan *Random Forest* untuk mendapatkan hasil prediksi harga tiket pesawat pada rute Jakarta-Surabaya dan Jakarta-Manado. Untuk setiap model prediksi akan dilakukan *parameter tuning* untuk mendapatkan model yang paling optimal. Setelah didapatkan tiga model yang optimal, dilakukan evaluasi model untuk menentukan model terbaik untuk setiap rute. Evaluasi model dilakukan dengan membandingkan nilai *Mean Absolute Error*, *Mean Absolute Percentage Error*, *Root Mean Squared Error*, dan Koefisien Determinasi dari tiap model. Hasil penelitian menunjukkan bahwa model yang memiliki akurasi paling baik untuk kedua rute tersebut adalah model *Random Forest*.

Kata Kunci : *regression tree, bagging regression tree, random forest, machine learning*

Referensi : 26 (1967-2021)

ABSTRACT

Kanaya Tabhita Djie (01112180013)

FLIGHT FARE IN INDONESIA PREDICTION USING TREE-BASED MODELS

Thesis, Faculty of Science and Technology (2023)

(xvi + 92 pages, 47 figures, 25 tables, 2 appendices)

Every day, the data generated from the aviation industry in Indonesia is enormous. Available data such as departure time, arrival time, number of transits, scheduled flight day, and so on may be used to predict flight fares. This study uses three machine learning techniques, namely Regression Tree, Bagging Regression Tree, and Random Forest to get the flight fares prediction result for Jakarta-Surabaya and Jakarta-Manado routes. For each prediction model, parameter tuning will be carried out to obtain the most optimal models. After obtaining three optimal models, model evaluation is performed to determine the best model for each route. Model evaluation is done by comparing each value of the Mean Absolute Error, Mean Absolute Percentage Error, Root Mean Squared Error and Coefficient of Determination of each model. The results reveal Random Forest as the model with the best accuracy for both routes.

Keywords : regression tree, bagging regression tree, random forest,
machine learning

Reference : 26 (1967-2021)