

ABSTRAK

PERBANDINGAN PERFORMANSI ALGORITMA *RANDOM FOREST*, KNN (*K-NEAREST NEIGHBORS*), DAN ADABOOST *REGRESSOR* PADA PERAMALAN HARGA BERLIAN

Oleh :

Fawzan Ibnu Fajar

1197050042

Selain banyak diminati dalam bentuk perhiasan, berlian merupakan salah satu materi terkeras yang pernah ada di muka bumi. Beberapa komunitas mulai melirik berlian untuk menjadi aset investasi mengingat nilainya yang luar biasa tinggi. Namun, permasalahan ditemukan ketika berlian tidak memiliki acuan harga layaknya aset investasi lain seperti emas ataupun saham. Perkembangan teknologi tentunya dapat mengatasi hal tersebut dengan implementasi algoritma *machine learning*. Algoritma *Random Forest*, *K-Nearest Neighbors*, dan *Adaptive Boosting* (Adaboost) *Regressor* dipilih menjadi algoritma yang akan dibandingkan. *Dataset* yang digunakan pada penelitian ini berasal dari *kaggle repository* milik vittoriogiatti dengan jumlah sebanyak 53.940 baris yang setelah dilakukan *pre-processing* berkurang menjadi 47.524 baris. Hasil dari pelatihan dan pengujian masing masing model ialah *Random Forest Regressor* memiliki rata rata nilai *error* 19,85 (MSE *train*), 132,73 (MSE *test*), 4,45 (RMSE *train*), 11,52 (RMSE *test*) serta rata rata durasi pelatihan yaitu 13,73 detik. Sementara rata-rata nilai R2-Score sebesar 0,981. Adapun algoritma KNN *Regressor* memiliki rata-rata nilai R2-Score sebesar 0,964. Sedangkan Adaboost *Regressor* dengan rata rata R2-Score sebesar 0.884. Hasil dari penelitian ini menunjukkan bahwa *Random Forest Regressor* keluar menjadi algoritma terbaik diantara dua lainnya.

Kata Kunci : Regresi, Prediksi Harga Berlian, *Random Forest*, *K-Nearest Neighbors*, *Adaptive Boosting* (Adaboost) *Regressor*.

ABSTRACT

PERFORMANCE COMPARISON OF RANDOM FOREST, KNN (K-NEAREST NEIGHBORS), AND ADABOOST REGRESSOR ALGORITHMS IN DIAMOND PRICE FORECASTING

By :

Fawzan Ibnu Fajar

1197050042

Besides being in great demand in the form of jewelry, diamonds are one of the hardest materials that have ever existed on earth. Some communities are starting to look at diamonds to become investment assets considering their extraordinary high value. However, problems are found when diamonds do not have a reference price like other investment assets such as gold or stocks. Technological developments can of course overcome this by implementing machine learning algorithms. The Random Forest, K-Nearest Neighbors, and Adaptive Boosting (Adaboost) Regressor algorithms were chosen as the algorithms to be compared. The dataset used in this study comes from Vittoriogiatti's Kaggle repository with a total of 53,940 lines. The results of this study indicate that the Random Forest Regressor comes out as the best algorithm among the other two, with an average error value of 19.85 (MSE train), 132.73 (MSE test), 4.45 (RMSE train), 11.52 (RMSE test) and the average training duration is 13.73 seconds. While the R2-Score value is 0.981 and produces predictions with a very small difference from the original value of 0.2. The KNN Regressor algorithm has an average R2-Score value of 0.964. While Adaboost Regressor with an average R2-Score of 0.884.

Keywords : *Regression, Diamond Price Prediction, Random Forest, K-Nearest Neighbors, Adaptive Boosting (Adaboost) Regressor.*