

ABSTRAK

Estimasi harga rumah merupakan permasalahan kompleks yang dipengaruhi oleh berbagai faktor seperti lokasi, luas bangunan, luas tanah, dan fasilitas pendukung. Penelitian ini bertujuan untuk menghasilkan estimasi harga rumah di wilayah Yogyakarta menggunakan kombinasi metode *Random Forest* dan *Extreme Gradient Boosting Tree (XGBoost)*. Data yang digunakan dikumpulkan melalui *web scraping* dari situs properti 99.co, yang setelah melalui tahap pembersihan menghasilkan 2.064 data. Proses penelitian meliputi tahapan *data understanding*, *preprocessing*, *modeling*, evaluasi, hingga *deployment* dalam bentuk sistem berbasis web. Model *Random Forest* dan *XGBoost* masing-masing dioptimasi melalui *hyperparameter tuning* untuk meningkatkan performa estimasi. Evaluasi dilakukan menggunakan metrik *Root Mean Squared Error (RMSE)*, *Mean Absolute Error (MAE)*, dan koefisien determinasi (R^2). Hasil pengujian menunjukkan bahwa kombinasi kedua model mampu memberikan estimasi harga rumah yang lebih stabil dan mendekati kondisi pasar aktual di Yogyakarta. Penelitian ini diharapkan dapat menjadi referensi berbasis data dalam pengambilan keputusan jual beli maupun investasi properti.

Kata Kunci: Estimasi harga rumah, *Random Forest*, *XGBoost*, *Machine Learning*, Yogyakarta.

ABSTRACT

House price estimation is a complex problem influenced by various factors such as location, building size, land area, and supporting facilities. This research aims to produce a house price estimation model for the Yogyakarta area using a combination of Random Forest and Extreme Gradient Boosting Tree (XGBoost) methods. The data were collected through web scraping from the 99.co property website, resulting in 2,064 cleaned entries after preprocessing. The research process includes data understanding, preprocessing, modeling, evaluation, and system deployment in the form of a web-based application. Both the Random Forest and XGBoost models were optimized through hyperparameter tuning to enhance estimation performance. Model evaluation was carried out using Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the coefficient of determination (R^2) metrics. The experimental results show that the combination of the two models provides more stable and accurate house price estimations that closely reflect actual market conditions in Yogyakarta. This research is expected to serve as a data-driven reference for decision-making in property purchasing and investment.

Keywords: *House Price Estimation, Random Forest, XGBoost, Machine Learning, Yogyakarta.*