HOUSE PRICE PREDICTION SYSTEM BASED ON FEATURES USING RANDOM FOREST ALGORITHM IN THE REGION OF YOGYAKARTA

MUHAMAD SYAHDAN RAMDHANI
Informatics Study Program, Faculty of Science and Technology
University of Technology Yogyakarta
Jl. Ringroad Utara, Jombor, Sleman, Yogyakarta
E-mail: msyahdanramdhani@gmail.com

ABSTRACT

This web-based house price prediction system for the Special Region of Yogyakarta was developed using the Random Forest Regression algorithm. The study addresses the inefficiencies of traditional house price estimation methods, which rely heavily on direct observation and visual assessments. Housing data were collected through web scraping from property marketplace websites and processed through the stages of Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation, and Deployment. The model utilizes property features such as location, land area, building area, number of bedrooms and bathrooms, and the presence of a carport. Model performance was evaluated using Root Mean Square Error (RMSE) and R-squared (R^2) metrics, resulting in an RMSE of 899,623,805.89 and an R^2 of 0.7376, indicating a fairly strong predictive capability. Predicted house prices are presented both numerically and through interactive visualizations to enhance user comprehension. This system is expected to assist prospective buyers and investors in making more objective and data-driven decisions.

***Keywords**: House Price Prediction, Random Forest, Web Scraping, Machine Learning*