Journal of Artificial Intelligence and Engineering Applications

Website: https://ioinformatic.org/

15th February 2025. Vol. 4. No. 2; e-ISSN: 2808-4519

House Price Prediction Analysis Using a Comparison of Machine Learning Algorithms in the Jabodetabek Area

Indah Ratna Ningsih[1*], Ahmad Faqih[2], Ade Rizki Rinaldi[3]

[1,2] *Teknik Informatika, STMIK IKMI Cirebon*
[3]*Rekayasa Perangkat Lunak, STMIK IKMI Cirebon*
indahratnan23@gmail.com*

Abstract

Jabodetabek, as the largest metropolitan area in Indonesia, has complex property price dynamics, making it difficult for developers and buyers to determine house prices. This study aims to analyze and compare the performance of the Multiple Linear Regression and Random Forest Regression algorithms in predicting house prices in the region. The data was obtained through scraping techniques from the rumah123.com website in October 2024, covering 999 data points with variables such as price, location, building area, land area, number of bedrooms, bathrooms, and garages. A comparative approach with cross-validation was applied to evaluate the performance of both algorithms using the metrics MAE, MSE, RMSE, MAPE, and R^2. The research results show that Random Forest Regression using GridsearchCV has better predictive performance, with an MAE value of Rp.645,764,815, MAPE of 28.12%, and R^2 of 0.864. The main factors influencing house prices in Jabodetabek include building size, land size, number of bedrooms, bathrooms, garages, and location. This finding emphasizes the superiority of Random Forest Regression in capturing complex data patterns and the significant role of these variables in determining house prices.

Keywords: *House Price; Multiple Linear Regression; Random Forest Regression; Machine Learning; Prediction*

1. Introduction

Jabodetabek is one of the largest metropolitan areas in Indonesia. Property prices in the Jabodetabek area are highly volatile due to urbanization, economic growth, and continuously increasing market demand. The property industry faces a major problem, especially in the Jabodetabek area, which is the difficulty in determining the right house price. House prices are influenced by a variety of complex factors, such as location, building size, land area, number of bedrooms, bathrooms, and other facilities. There are significant price differences in various areas as well as rapid market fluctuations. Therefore, property developers and buyers face a significant challenge in making the right decisions about house prices. The comparison of machine learning algorithms for house price prediction shows that each algorithm has its own advantages, depending on the complexity of the dataset and the variables to be predicted [1]. These results emphasize the importance of choosing the right loss function for machine learning models to ensure high prediction accuracy and model reliability, especially when predicting property prices. Loss functions play an important role in measuring how well the model makes predictions, and choosing an appropriate one can minimize prediction errors. In the context of property price prediction, where the factors affecting prices are highly variable and complex, using the right loss function will ensure the model is not only accurate, but also reliable to provide consistent and realistic predictions [2].

Using a large dataset with complex and non-linear relationships, Random Forest has demonstrated its potential in predicting volatile market prices, such as those of Bitcoin [3]. All of this indicates that algorithm selection should be based on data characteristics, with Random Forest being the best choice for datasets that allow for complex interactions between elements [4]. These results highlight the importance of selecting the right loss function for machine learning models to ensure high prediction accuracy and model reliability, especially in predicting property prices. Loss functions play an important role in measuring the extent to which a model can produce accurate predictions, and by selecting appropriate functions, we can minimize prediction errors. In the context of property price prediction, where the factors affecting prices are highly variable and complex, the use of an appropriate loss function will ensure the model is not only accurate, but can also be relied upon to provide consistent and realistic predictions. Thus, choosing the right loss function is not just a matter of improving prediction results, but also ensuring that the model can perform stably under various dynamic property market conditions [5]. These results emphasize the importance of selecting an appropriate loss function for machine learning models in order to produce high prediction accuracy and ensure model reliability, especially in predicting property prices. Loss functions play an important role in measuring the extent to which a model can produce accurate predictions, and by choosing the right function, we can minimize the prediction error. In the context of property price prediction, which is influenced by a wide range of highly variable and complex factors, the use of an appropriate loss function will ensure that the model is not only accurate, but can also be relied upon to provide consistent and realistic predictions. Therefore, choosing an appropriate loss function is not only to improve the prediction results, but also to ensure that the model can perform

stably across a wide range of dynamic property market conditions [6]. Linear regression is a statistical method used for analysis in various fields, including financial analysis, where it is used to predict stock prices [7].

In predictive analysis, linear regression remains a key technique, especially when studying the relationship between independent and dependent variables in structured datasets [8]. A popular method for predicting continuous variables, like those used in mobile second price prediction, is Multiple Linear Regression , which successfully determines the link between independent and target variables [9]. The Linear Regression Algorithm has proven to be effective in predicting patent prices in Indonesia, analyzing its performance in predicting results based on historical data [10]. All of this suggests that the efficacy of predictive modeling in practical statistical analysis is greatly influenced by data analysis employing EDA and fitter key selection [11]. For the Random Forest model to be optimized for predictive accuracy in real estate applications, feature selection and hyperparameter tweaking are essential [12]. As evidenced by the prediction of the Water Quality Index (WQI), where fine-tuning produces more accurate and dependable forecasts, the use of hyperparameter tuning in regression algorithms improves the performance of predictive models [13]. Because it enables the model to better fit the data, hyperparameter tuning in regression models has been shown to increase prediction accuracy in the context of automobile price prediction [14]. To thoroughly evaluate each algorithm's efficacy in predicting home values, evaluation metrics like MAE, MAPE, RMSE, and R^2 are employed [15]. This research aims to determine house price predictions and analyze the factors influencing them, as well as to identify which algorithm is more accurate in predicting house prices in the Jabodetabek area.

2. Research Methodology

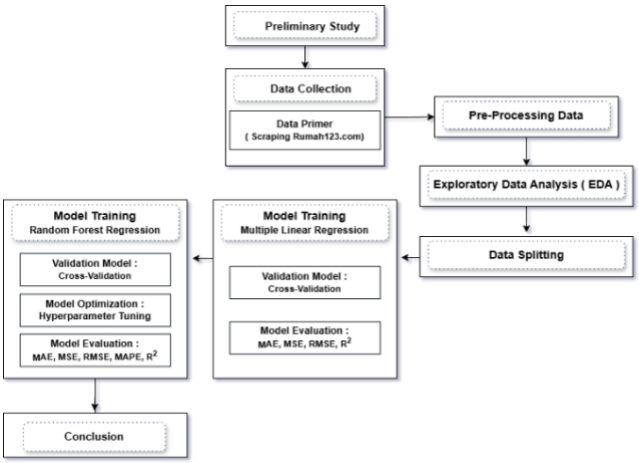

Fig. 1: Research methodology

Fig. 1 illustrates the research flow based on the Knowledge Discovery in Databases (KDD) approach, designed to systematically predict house prices.

2.1. Preliminary Study

Determining the challenges, objectives, and benefits of a study is an important first step in understanding the topic being researched. This process helps in identifying the key aspects that need to be analyzed more deeply. In order to find the best solution to the problem at hand, literature research was conducted by studying relevant algorithms, such as Random Forest Regression and Multiple Linear Regression. By studying these algorithms, we can understand the strengths and weaknesses of each, as well as determine the most appropriate method to apply to the problem being studied, so that the research can run effectively and provide optimal results.

2.2. Data Collection

Primary data from the web scraping approach is employed in this study's data collection methodology. Web scraping is the technique of autonomously obtaining structured data from web pages. In this instance, pertinent information was extracted from the Rumah123.com website in October 2024, including the house ‘Price’, ‘Building Size’, ‘Land Size’, number of ‘Bedrooms’, number of ‘Bathrooms’, ‘Garage’, and ‘Location’. The goal of the scraping procedure is to gather data from property records that are pertinent to the goals of the study. This makes it possible for researchers to gather a lot of information on different places and kinds of features.

2.3. Pre-processing Data

The preprocessing step is subsequently applied to the gathered data. This involves handling outliers, imputation to fill in missing values, data normalization or modification when necessary, and data cleaning to get rid of duplicates.

2.4. Exploratory Data Analysis (EDA)

Before creating a predictive model, do an exploratory investigation of the data to learn about its structure, properties, and trends as well as how it is distributed. Descriptive analysis and visualization are the main topics of this stage.

2.5. Data Splitting

The procedure of data splitting is used to divide the data into two subsets: 20% for testing and 80% for training. Random Forest Regression and Multiple Linear Regression methods are trained on the training data, and the model's ability to forecast home prices is assessed using the test data. Twenty percent of the data is used for testing, while eighty percent is used for training.

2.6. Model Training

The goals of this study's model training is to create a home price prediction model using previously processed data. Random Forest Regression and Multiple Linear Regression are the two primary methods employed in this study.

The following actions were taken for every model:

a. Validation of the model with when measuring the model's overall performance using data that was not observed during training, cross-validation is used. Several folds are created from the training data using this method, and each fold is utilized as validation data in turn.
b. An optimization model that uses hyperparameter tuning to enhance Random Forest Regression's performance. This procedure guarantees that the model performs best on the dataset that is being utilized.

2.7. Model Evaluation

Metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), and R-squared (R^2) are used to assess the performance of both models.

2.8. Conclusion

Based on the evaluation results, conclusions were drawn to determine the superior algorithm for predicting house prices in the Jabodetabek area. And visualize the results between the predicted values and the actual values for both models.

3. Results And Discussion

3.1 Data Collection

The data collected in this research was gathered through web scraping techniques from the Rumah123.com website. The data was collected in October 2024 and focused on the Jabodetabek area. The information obtained includes Location, Land Size, Building Size, Bedrooms, Bathrooms, Garage, and Price.

3.2 Import Library

In the first stage, various necessary Python libraries were imported to support data processing, Exploratory Data Analysis (EDA), modeling, and deployment. Each library has specific functions that help streamline model analysis and implementation. The following are the main functions of the libraries used:

1. Pandas is a very popular library for analyzing and manipulating data in tabular formats, such as CSV or Excel files. It offers powerful tools for handling large amounts of data, allowing users to filter, combine and summarize data with ease.
2. NumPy is a library designed to perform complex mathematical operations, including linear algebra and multidimensional matrix and array manipulation. It is often the basis for other libraries that require numerical computation.
3. Matplotlib provides the ability to create various types of graphs and data visualizations. From line graphs, bar charts, to scatter plots, this library allows users to clearly display data patterns and trends.
4. Seaborn equipped Matplotlib with a simpler interface and more appealing aesthetics, making it a top choice for creating statistical visualizations such as heatmaps, box plots, and information-rich scatter plots.
5. Scikit-learn (sklearn) is the core library for machine learning in Python. With various algorithms available, this library facilitates modeling, evaluation, and optimization of models. Its features include regression, classification, clustering, and evaluation techniques such as cross-validation.
6. SciPy is a versatile library for scientific and technical computing. With modules that support linear algebra, statistics, optimization and more, SciPy is an essential tool for in-depth analysis.

The collaboration of these libraries provides a solid framework for the entire process, from raw data processing to predictive model implementation, enabling better data-driven decision making.

3.3 Data Pre-Processing

Data preprocessing is one of the important stages that serves to ensure the quality of the data used in the model. At this stage, various steps are taken to clean and prepare the data for use in the analysis. The data cleaning process is used to remove unnecessary data duplication, while outlier handling is done to address extreme values that may affect the analysis results. In addition, imputation is used to fill in missing values, so that the data becomes complete and no information is missed. Finally, data normalization or transformation steps are applied as

needed, so that the data has an appropriate scale and can improve prediction accuracy. All of these steps are important to ensure that the data used can provide optimal and reliable results in modeling.

Table 1: Data Before Pre-processing

No	Location	Land Size	Building Size	Bedrooms	Bathrooms	Garage	Price
1	Jakarta Selatan	866	684	6.0	3.0	14.0	6.000000e+10
2	Jakarta Selatan	600	500	4.0	4.0	NaN	1.690000e+10
3	Jakarta Selatan	53	130	4.0	4.0	1.0	1.750000e+09
4	Jakarta Selatan	684	400	4.0	3.0	NaN	2.800000e+10
5	Jakarta Selatan	150	250	5.0	3.0	1.0	4.250000e+09

Because Table 1 has missing data and incredibly high values, the data must be pre-processed to fix these problems, producing data similar to Table 2.

Table 2: Data After Pre-processing

No	Location	Land Size	Building Size	Bedrooms	Bathrooms	Garage	Price
1	Jakarta Selatan	336.0	427.0	5.0	3.0	3.0	6.000000e+10
2	Jakarta Selatan	336.0	427.0	4.0	4.0	2.0	1.690000e+10
3	Jakarta Selatan	53.0	130.0	4.0	4.0	1.0	1.750000e+09
4	Jakarta Selatan	336.0	400.0	4.0	3.0	2.0	2.800000e+10
5	Jakarta Selatan	150.0	250.0	5.0	3.0	1.0	4.250000e+09

Table 3: Data Location After One-Hot Encoding

No	Bekasi	Bogor	Depok	Jakarta Barat	Jakarta Pusat	Jakarta Selatan	Jakarta Timur	Jakarta Utara	Tangerang	Tangerang Selatan
1	False	False	False	False	False	True	False	False	False	False
2	False	False	False	False	False	True	False	False	False	False
3	False	False	False	False	False	True	False	False	False	False
4	False	False	False	False	False	True	False	False	False	False
5	False	False	False	False	False	True	False	False	False	False

Table 3 shows the results of the one-hot encoding process on the location variable to convert categorical data into a numerical format that can be used in modeling. In this process, each unique category of the location variable is represented as a separate column, where a value of 1 indicates the presence of that location in the data, while a value of 0 indicates its absence. This approach allows machine learning algorithms, which work with numerical data, to make optimal use of location category information without any bias regarding the order or hierarchy between categories. As such, location variables can contribute significantly to improving model accuracy in analysis and prediction.

3.4 Exploratory Data Analysis (EDA)

Before entering the model training stage, the Exploratory Data Analysis (EDA) phase aims to deeply understand the characteristics of the data to be used. This analysis involves identifying patterns, distributions, and relationships between variables, as well as detecting problems that may exist in the data, such as blank values, outliers, or inconsistencies. With EDA, important insights can be gained, such as which variables have a significant influence, relationships between variables that need attention, and trends or anomalies that may affect the results of the analysis. This process not only helps in making decisions about data preprocessing but also ensures that the data is optimally ready to be used for model training.

Table 4: Descriptive Data Analysis

	Land Size	Building Size	Bedrooms	Bathrooms	Garage	Price
Count	999.000000	999.000000	999.000000	999.000000	999.000000	999.0000+02
Mean	152.736737	182.628629	3.338338	2.631632	1.797798	3.345133e+09
Std	97.151029	130.144259	0.957850	0.955126	0.565504	2.721676e+09
Min	16.000000	1.000000	2.000000	1.000000	1.000000	1.750000e+08
25%	80.000000	80.000000	3.000000	2.000000	1.000000	114.0000e+09
50%	120.000000	131.000000	3.000000	3.000000	2.000000	2.350000e+09
75%	208.000000	253.500000	4.000000	3.000000	2.000000	4.830000e+09
Max	336.000000	427.000000	5.000000	4.000000	3.000000	8.520000e+09

In Table 4, the values of the data can be seen, starting from total (count), standard deviation (std), mean, minimum (min), quartiles (25%, 50%, 75%), and maximum values (max).

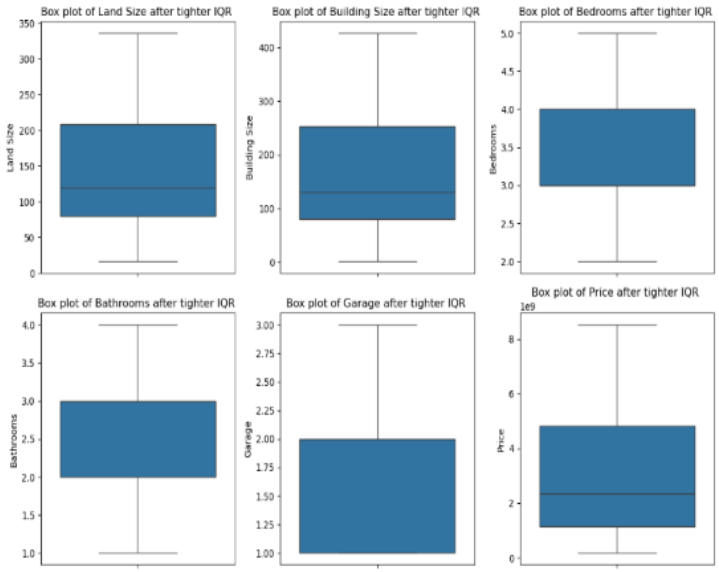

Fig. 2: Boxplot Data Distribution

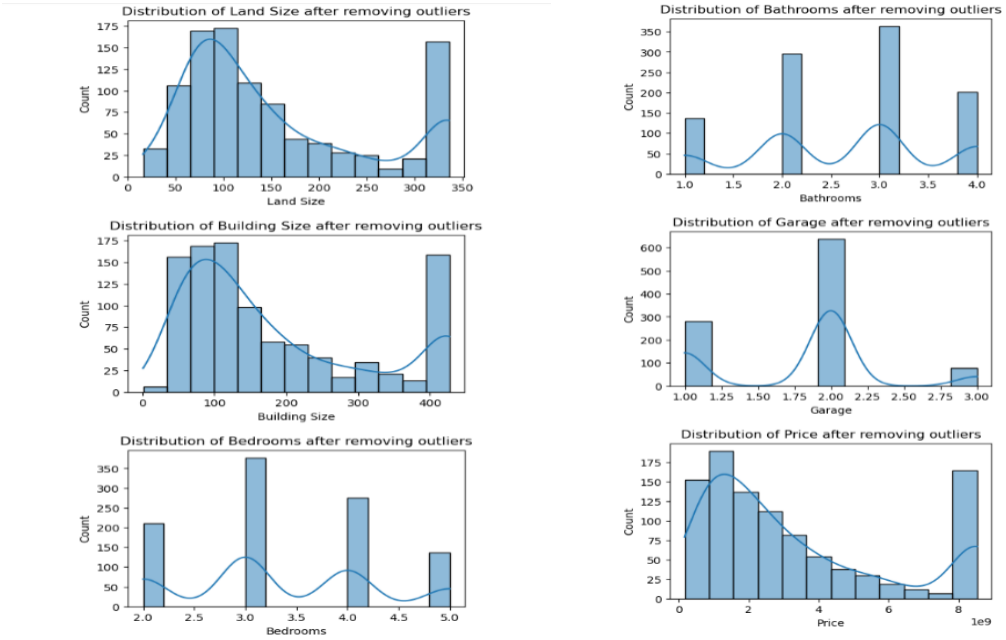

Fig. 3: Displot Data Distribution Pattern

In Fig. 2 and 3, the distribution of data on house 'Price', 'Building Size', 'Land Size', number of 'Bedrooms', number of 'Bathrooms', 'Garage', and 'Location' that have undergone pre-processing can be seen.

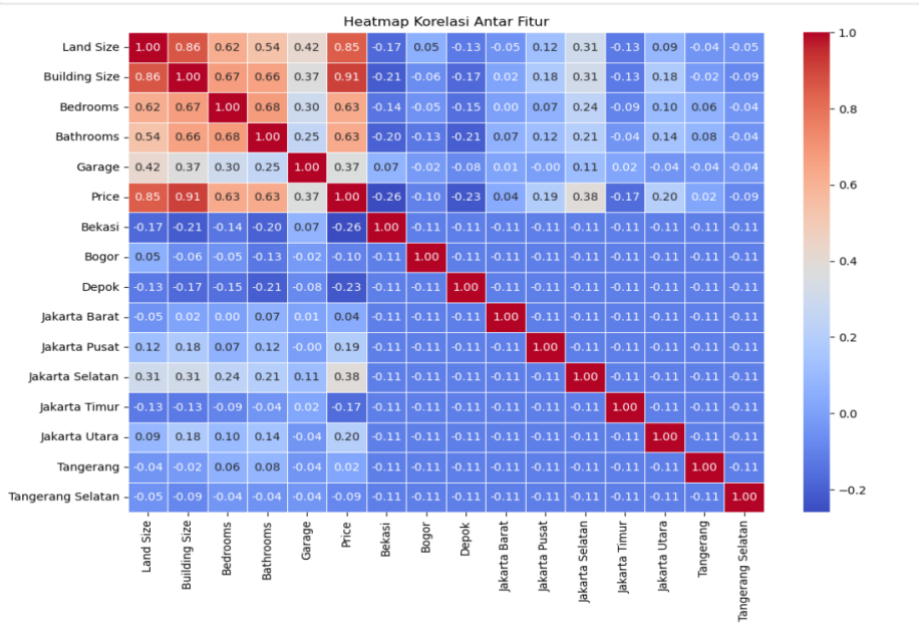

Fig. 4: Correlation Heatmap Between Variables

The correlation heatmap above show that, with correlation values of 0.85, 0.91, 0.63, and 0.63 respectively, the features of Land Size, Building Size, Bedrooms, and Bathrooms have a high positive correlation with house prices. This shows that a home's price increases with the size of the land, the building, and the number of bedrooms and bathrooms. However, because they are encoding variables that only show particular regional categories, fake location features like Bekasi, Bogor, and others have little link with price and other features. All things considered, property size has a bigger influence on home values than location.

3.5 Data Splitting

The dataset of 999 data points was divided into two subsets, the training set and the testing set, with a ratio of 80:20. A total of 799 data points were used in the training set to train the model, while the remaining 200 data points were included in the testing set to evaluate the performance of the model against data that had never been seen before. The distribution was randomized to ensure that both subsets had a representative distribution of data, so that the model could be tested objectively and reliably. This approach aims to reduce potential bias and ensure that the model is able to capture patterns from the training data while providing good prediction performance on new data.

3.6 Model Training Multiple Linear Regression

In this analysis, the sklearn library was used to calculate Multiple Linear Regression, with house price as the variable to be predicted (y) and factors such as building size, land size, number of bedrooms, bathrooms, and garages as influencing factors (x). This regression model is trained using training data to find a linear relationship between the house price and these factors. During training, regression coefficients are calculated to show how much influence each factor has on house prices. To ensure the model can perform well on new data, cross-validation is performed, which helps evaluate how well the model can generalize its results without overfitting.

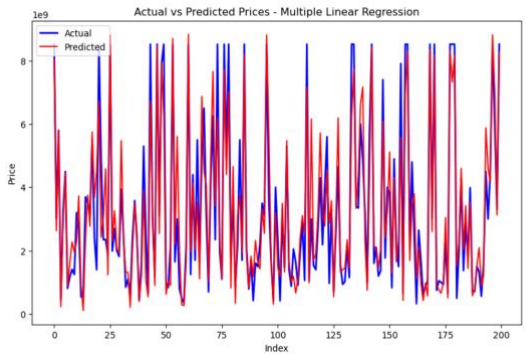

Fig. 5: Line Plot Multiple Linear Regression

The actual and predicted values of house prices generated by the Multiple Linear Regression model are compared in Figure 5. This graph is used to assess the performance of the model by showing the extent to which the predictions generated by the model match the original data. By looking at the comparison between the actual house prices and the predicted prices, we can evaluate the accuracy of the model in predicting house prices and how well the model can capture patterns in the data.

3.7 Model Training Random Forest Regression

The Random Forest Regression model was trained using the training data (80% of the dataset, i.e. 799 data points) to study the relationship between the independent and dependent variables. The independent variable (x) includes columns such as Land Area, Building Area, Number of Bedrooms, Number of Bathrooms, and Garage, while the dependent variable (y) is the Price column. With the ensemble principle, the Random Forest model combines predictions from various Decision Trees to produce more accurate results. In order for the model to generalize well, the Cross-Validation technique is used to divide the training data into several subsets, which are then used alternately to train and validate the model. Afterwards, GridSearchCV and RandomSearchCV are used for Hyperparameter Tuning, which aims to find the best combination of parameters to improve the model's performance, ensuring that the model functions optimally under various conditions.

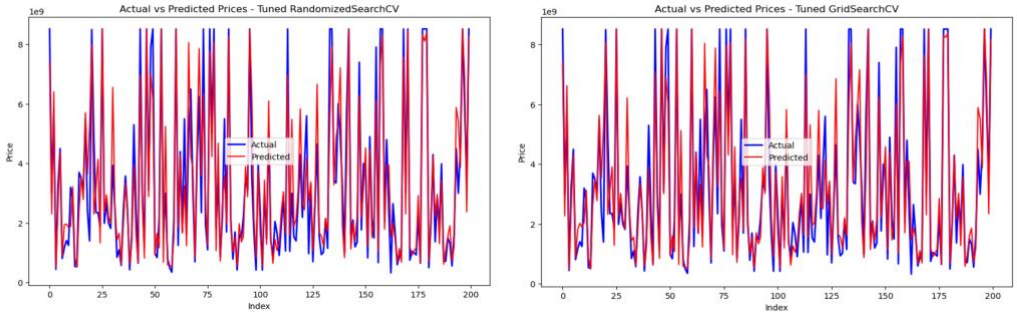

Fig. 6: Line Plot Random Forest Regression

The actual values and house prices predicted by the Random Forest Regression model that has gone through the GridSearchCV and RandomSearchCV processes are shown in Figure 6. This graph is used to evaluate the performance of the model by comparing the extent to which the predictions generated by the model match the original data. By looking at this comparison, we can assess how accurate the model is in predicting house prices and how well it captures the patterns in the data, giving us an idea of the effectiveness of the model's predictions against the actual values.

3.8 Evaluation Model

Various metrics are used to evaluate the performance of prediction models, including Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), and R-squared (R^2). MAE measures the average absolute error magnitude between predicted and actual values, giving an idea of the average degree of deviation without taking into account the direction of the error. MSE calculates the average square of the error, which penalizes larger errors and is therefore more sensitive to outliers. RMSE, as the square root of MSE, returns the error scale to the original units of the data, easing the interpretation of the results. MAPE presents the average error in percentage form, providing an understanding of the proportion of error relative to the actual value. Meanwhile, R^2 measures the extent to which the model is able to explain the variability of the data, with values close to 1 indicating excellent model performance. The combination of these metrics allows for a comprehensive analysis, providing deep insight into the strengths and weaknesses of each model.

Table 5: Model Evaluation Results

Model	MAE	MSE	RMSE	MAPE	R^2
Multiple Linear Regression	691,484,839	8.91×10^{17}	943,681,116	32 %	0,864
Random Forest Regression (GridSearchCV)	645,764,815	8.96×10^{17}	946,430,013	28 %	0,863
Random Forest Regression (RandomizedSearchCV)	668,714,344	9.44×10^{17}	971,422,791	29 %	0,856

4. Conclusion

This research successfully shows that the Random Forest Regression algorithm has superior performance compared to Multiple Linear Regression in predicting house prices in the Jabodetabek area. With a Mean Absolute Error (MAE) value of Rp645,764,815, a Mean Absolute Percentage Error (MAPE) of 28.12%, and a coefficient of determination (R^2) of 0.864, Random Forest Regression is able to capture complex data patterns with high accuracy. The main factors that influence house prices include building area, land area, number of bedrooms, number of bathrooms, presence of a garage, and property location. The findings confirm the superiority of the Random Forest Regression algorithm in analyzing dynamic property data, providing valuable insights for developers and buyers in accurately determining house prices. In addition, the results of this study can be an important reference to support strategic decision-making in the property industry.

References

[1] C. Haryanto, N. Rahaningsih, and F. M. Basysyar, "Komparasi Algoritma Machine Learning Dalam Memprediksi Harga Rumah," *Jurnal Mahasiswa Teknik Informatika (JATI)*, vol. 7, no. 1, pp. 533–539, Feb. 2023.

[2] A. Hjort, J. Pensar, I. Scheel, and D. E. Sommervoll, "House Price Prediction With Gradient Boosted Trees Under Different Loss Functions," *Journal of Property Research*, vol. 39, no. 4, pp. 338–364, 2022, doi: 10.1080/09599916.2022.2070525.

[3] S. Saadah and H. Salsabila, "Jurnal Politeknik Caltex Riau Prediksi Harga Bitcoin Menggunakan Metode Random Forest (Studi Kasus: Data Acak Pada Awal Masa Pandemic Covid-19)," May 2021. [Online]. Available: https://jurnal.pcr.ac.id/index.php/jkt/

[4] S. Fachid and A. Triayudi, "Perbandingan Algoritma Regresi Linier dan Regresi Random Forest Dalam Memprediksi Kasus Positif Covid-19," *Jurnal Media Informatika Budidarma*, vol. 6, no. 1, p. 68, Jan. 2022, doi: 10.30865/mib.v6i1.3492.

[5] I. M. Mufidah, H. Basuki, P. Ilmu, and K. Masyarakat, "Analisis Regresi Linier Berganda Untuk Mengetahui Faktor Yang Mempengaruhi Kejadian Stunting Di Jawa Timur," *Indonesian Nursing Journal of Education and Clinic*, vol. 3, no. 3, pp. 51–59, Mar. 2023.

[6] D. Alita, A. D. Putra, and D. Darwis, "Analysis Of Classic Assumption Test And Multiple Linear Regression Coefficient Test For Employee Structural Office Recommendation," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 15, no. 3, p. 295, Jul. 2021, doi: 10.22146/ijccs.65586.

[7] L. Alpianto and A. Hermawan, "Moving Average untuk Prediksi Harga Saham dengan Linear Regression," *Jurnal Buana Informatika*, vol. 14, pp. 117–126, Nov. 2023.

[8] S. Hanifah, F. Akbar, and R. P. Santi, "Implementasi Business Intelligence dan Prediksi Menggunakan Regresi Linear pada Data Penjualan dan Breakage di PT XYZ," *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 8, no. 3, pp. 144–152, Dec. 2022, doi: 10.25077/teknosi.v8i3.2022.144-152.

[9] D. M. Huda, G. Dwilestari, and A. R. Rinaldi, "Prediksi Harga Mobil Bekas Menggunakan Algoritma Regresi Linear Berganda," *Jurnal Informatika dan Rekayasa Perangkat Lunak*, vol. 6, no. 1, pp. 150–157, Mar. 2024.

[10] D. Novianty, N. D. Palasara, and M. Qomaruddin, "Algoritma Regresi Linear pada Prediksi Permohonan Paten yang Terdaftar di Indonesia," *Jurnal Sistem dan Teknologi Informasi (Justin)*, vol. 9, no. 2, p. 81, Apr. 2021, doi: 10.26418/justin.v9i2.43664.

[11] F. M. Basysyar and G. Dwilestari, "House Price Prediction Using Exploratory Data Analysis and Machine Learning with Feature Selection," *Acadlore Transactions on AI and Machine Learning*, vol. 1, no. 1, pp. 11–21, Nov. 2022, doi: 10.56578/ataiml010103.

[12] A. B. Adetunji, O. N. Akande, F. A. Ajala, O. Oyewo, Y. F. Akande, and G. Oluwadara, "House Price Prediction using Random Forest Machine Learning Technique," in *Procedia Computer Science*, Elsevier B.V., 2022, pp. 806–813. doi: 10.1016/j.procs.2022.01.100.

[13] M. Radhi, S. H. Sinurat, D. R. H. Sitompul, E. Indra, and S. Informasi, "Prediksi Water Quality Index (Wqi) Menggunakan Algoritma Regressi Dengan Hyper-Parameter Tuning," *Jurnal Sistem Informasi dan Ilmu Komputer Prima)*, vol. 5, no. 1, pp. 44–50, Aug. 2021.

[14] M. Radhi, D. R. H. Sitompul, S. H. Sinurat, and E. Indra, "Prediksi Harga Mobil Menggunakan Algoritma Regressi Dengan Hyper-Parameter Tuning," *Jurnal Sistem Informasi dan Ilmu Komputer Prima*, vol. 4, no. 2, pp. 28–32, Feb. 2021.

[15] E. Fitri, "Analisis Perbandingan Metode Regresi Linier, Random Forest Regression dan Gradient Boosted Trees Regression Method untuk Prediksi Harga Rumah," *Journal Of Applied Computer Science And Technology(JACOST)*, vol. 4, no. 1, pp. 2723–1453, 2023, doi: 10.52158/jacost.491.