

JURNAL TEKNOBISNIS 2023 Vol. 9(1), 49-59

Interdisciplinary School of Management and Technology Institut Teknologi Sepuluh Nopember Received 26 Feb, 2023; Revised 9 Apr, 2023; Accepted 30 Apr, 2023 | DOI: 10.12962/j24609463.v9i1.930

# Prediction of Surabaya City Apartment Rental Price Using the Ensemble Method

Erlina Komaruljannah<sup>1,2\*</sup>, Chastine Fatichah<sup>3</sup>

# ABSTRACT

Increasing the need for residential space requires in city to improve development infrastructure to meet the needs of its inhabitants. Having an apartment will help optimize the capacity of residential land in limited land, especially for areas with high population density and small areas such as the city of Surabaya. Until 2020, the level of demand for housing or rental in apartments tends to decline and after the COVID-19 pandemic made a huge impact on the apartment market from a sales perspective and prices dropped dramatically. A property consulting firm, Coldwell Banker Commercial, stated that among other big cities, only Surabaya had a better performance in apartment marketing, because the sales rate and average price increased by 0.4%. Seeing the urgency and potential for the weakening of modern society in choosing apartment housing in the city of Surabaya, the authors conducted a study that could assist in predicting apartment rental prices according to apartment criteria in Surabaya such as sub-district location, unit type, area, floor number, and unit facility contents. The data used is data from travelio.com which is one of the most trusted websites in Indonesia for apartment rentals using the scraping method. The apartment rental price prediction model using the ensemble method found that the stacking model with a combination of gradient boosting regression, random forest, decision tree regression, linear regression meta liner regression has higher performance, with an MAE value of IDR 243,401, RMSE of IDR 179,432 and a value R2 is 63.28%, which means that it affects the price of apartment rentals by 63.28%, while 36.72% is influenced by other factors that are not included in the model.

KEYWORDS: Rent Prices, Apartments, Bagging, Boosting, Stacking, Ensemble Learning

<sup>&</sup>lt;sup>1</sup>Staff, Public Works, Highways, and Drainage Department, Surabaya, Indonesia

<sup>&</sup>lt;sup>2</sup>Business Analytics, Interdisciplinary School of Management and Technology, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

<sup>&</sup>lt;sup>3</sup>Department of Informatics Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

<sup>\*</sup>Corresponding author: 6032211230@mhs.its.ac.id

#### **1. INTRODUCTION**

Surabaya City is the second metropolitan city after DKI Jakarta, experiencing rapid growth and development. With a very dense population, the development of Surabaya City is directed to become a regional growth center in East Java Province. This is influenced by population density and a reasonably high flow of migrants. The increase in development activities with various types of business activities in the form of industry, hospitals, hotels, malls, apartments, and offices in different cities of Surabaya is currently triggering urbanization (Sri Duto Wisnugroho, 2021).

Increasing the need for residential space requires a city to improve its development infrastructure to meet the needs of its inhabitants. Regarding the spatial pattern of the development of residential areas in Surabaya, namely vertical settlements in the form of flats (simple) and apartments or condominiums scattered in almost all corners of the city of Surabaya. In contrast, residential areas are directed to develop to the west, east, and south of the city (*Rencana Kerja Pemerintah Daerah Kota Surabaya Tahun 2022*, 2022) Based on data recorded by the Surabaya City Population and Civil Registry Office, administratively, the Surabaya City government consists of 31 sub-districts, 153 sub-districts, 1,360 Residential Units (RW) and 9,126 Neighborhood Units (RT). The registered population of Surabaya City as of December 2022 is 2,987,863 people, with a population density of 8,612 people per km2 (*Badan Pusat Statistik Kota Surabaya*, 2022). The population of Surabaya will continue to grow due to several factors, one of which is the emergence of new jobs in the trade and industrial sectors.

Overcome the limited amount of land in Surabaya and the large amount of land needed for the house, and this has led to efforts by the Government of the City of Surabaya to fulfill housing needs through the construction of housing with a vertical residential system, one of which is apartment housing. Apartments are synonymous with exclusive occupancy for certain groups that function to control land requirements and make space efficient. Coupled with the increasingly expensive land prices in the center of Surabaya City, building multi-story houses or apartments is one of the best solutions to overcome this problem (Angelia Gaina & Hariyanto, 2014). The apartment is a practical modern residence because the building is made vertically and very helpful in optimizing the capacity of occupants in limited land. The Ministry of Public Works and Public Housing (PUPR) is pushing for vertical residential development as part of the Transit Oriented Development (TOD) concept, namely residential areas integrated with public transportation nodes. Anita Firmanti as Plt. The Secretary General of the PUPR Ministry said, "Vertical residential developments such as flats, apartments, or condominiums will encourage more effective land use and a large number of residential units. In addition, by living in vertical housing in urban areas, millennials can more easily reach office and commercial areas where they work."

The growth of tall buildings in Surabaya during the last ten years has increased massively. One type of high-rise building that has the potential to be built in the city of Surabaya is an apartment or condominium. The construction of apartments is a form of

effort made by the Surabaya City Government to meet housing (Peraturan Daerah Kota Surabaya Nomor 12 Tahun 2014 Pasal 16, 2014). Surabaya City Apartments were built for the first time in 1985, namely Graha Residence Serviced, located in West Surabaya. Until 2022, 46 apartments were built across the north, east, central, west, and south areas (Department of Public Housing and Residential Areas and Land Affairs of Surabaya, 2022). Even though there is no specific plan for the designation of apartment land in Surabaya City planning, the Surabaya City Government continues to open vast opportunities in terms of apartment construction and development; this can be seen in the alternative settlement planning set out in the 2014-2034 Surabaya City Spatial Plan. In the 2014-2034 Surabaya City Spatial Planning, the Surabaya City Government provided great opportunities in the construction and development of apartments. This can be seen from the extensive distribution of apartment units built until 2016 of 19,905 units (Agmelina & Ariastita, 2017). The shift in interest in housing from landed houses to vertical houses is heavily influenced by today's people's lifestyles. In addition to the increasingly high prices of landed houses due to limited land, the lifestyle of urban communities that prioritizes efficiency, both in terms of time, distance, and energy, is a consideration in choosing vertical housing, most of which are built in big cities close to community centers or the Central Business District (CBD) (Krisnaputri, 2016)

Until 2020, the demand for housing or rental in apartments declined. After the COVID-19 pandemic hugely impacted the apartment market from a sales perspective, prices dropped dramatically. Property consulting company Coldwell Banker Commercial stated that the demand trend in 2021, when compared between guarter one and guarter two, demand for apartments in Surabaya fell 32%, Bandung 90%, Medan fell 100%, Semarang 100%, Batam 100%, Makassar 19%, Balikpapan stagnated, Bali fell 100%, while Palembang stagnated. Meanwhile, cities that experienced increased sales were only Surabaya, which was 0.4%, Bandung at 0.1%, and Makassar at 0.4%, the remaining 0%. Compared to other big cities, only Surabaya has a better marketing performance. Because of the sales level and average price increase in other big cities only, Surabaya has a better performance in apartment marketing because of the sales level and average price increase. Seeing the problems and potential weakening of modern society in choosing apartment dwellings in the city of Surabaya, the authors conducted a study that could assist in predicting apartment rental prices according to the classification of apartments in Surabaya, such as sub-district location, unit type, area, floor number, and unit facility contents.

To predict apartment rental prices in this study, use the ensemble method in supervised learning because the ensemble method is more robust and stable than other models. The Ensemble method is used to compare each algorithm for predicting apartment prices (Mulyahati, 2020). Research purposes of this research aims to analyzing data characteristics of apartments in Surabaya based on the travelio.com website, knowing the best performance evaluation of the prediction model to obtain an ensemble predictive model for apartment rental prices in Surabaya. This research will make it easier for prospective apartment tenants to be able to find apartment prices that follow their

financial capabilities, as well as apartment investors who buy apartments to be rented out again to be able to set rental prices according to the criteria for other apartment units. The benefits expected from this research are that the process and results described can be helpful to facilitate real estate marketing and apartment seekers in determining apartment rental prices that follow market prices and available apartment criteria, so that apartment sales can increase, mainly to deal with the problem of optimizing residential land in Indonesia.

# **2. LITERATURE REVIEW**

Apartment is a multi-story building built in an environment and divided into functionally structured sections in the horizontal and vertical directions which are units, each of which can be owned and used separately, especially for dwellings equipped with shared parts, shared objects, and common land (Angelia Gaina & Hariyanto, 2014) There is apartement classification, such as:

#### **Classification Based on Sleeping Space Per Occupational Unit**

The number of bedrooms per residential unit distinguishes apartments. An efficient apartment (efficiency apartment) is a residential unit with a main room used for various purposes. Sometimes this type is called a Studio Apartment or One Bedroom ( $\pm$  18.58 – 46.45 m2). In this type, the dining and sitting rooms are one; there is also a bedroom, kitchen, and bathroom/WC. Two-Bedroom Apartment, an apartment with two bedrooms is an apartment consisting of 2 bedrooms, a sitting room, a dining room, a kitchen, and a bathroom ( $\pm$ 46.45 - 92.90 m2). Three- Bedroom Apartment, an apartment with three bedrooms consists of 3 bedrooms, a sitting room, a dining room, a kitchen, and 1-2 bathrooms (55.74—111.48m2). Four-Bedroom Apartment, an apartment with four bedrooms consists of 4 bedrooms, a sitting room, a dining room, a kitchen, two bathrooms, and a warehouse with an area of ( $\pm$ 102.19-139.35m2) Five Bedroom Apartment, an apartment residence consisting of 5 bedrooms, a dining room, a sitting room, a sitting room, a sitting room, a sitting room, a kitchen, two bathrooms, and a warehouse with an area of ( $\pm$ 102.19-139.35m2) Five Bedroom Apartment, an apartment with five bedrooms, which can be called (a penthouse), is a luxury apartment residence consisting of 5 bedrooms, a dining room, a sitting room, a sitting room, a kitchen (complete with a pantry), three bathrooms with a dressing room, a service room, a laundry room, and storage (Kim & Kim, 2015)

#### **Classification Based on Type of Lease**

Apartments can be divided into types of rent among others: Ordinary rent when occupants can pay rent according to the agreement with the owner of the building/apartment residential unit without being bound by a time limit. Rent buy, rent purchase is the occupant pays rent which is used as installment purchases. The apartment residential unit will belong to the occupant if the installment reaches a predetermined price. Contract Lease, Contract rental is rent paid periodically by occupants to the owner of the building/apartment residential unit with periodic approval. If the contract period has ended, a new agreement can be entered into, or a new contract can be continued (Purnamasari et al., 2018).

#### **Classification Based on Services and Facilities**

Apartments can be differentiated based on their supporting facilities. Apartment Fully Service, an apartment that provides standard hotel services for its residents, such as laundry, catering, cleaning, etc. Apartment Fully Furnished, apartments that provide furniture or equipment in apartment units. Apartment Fully Furnished and Fully Service, a combination of the two types of apartments listed above. Apartment Building Only, apartments that do not provide room service and furniture (De Chiara & M. Crosbie, 2001)

#### **Classification Based on Accessibility**

Related to location, one of the factors that determine whether a location is attractive is the level of accessibility. The accessibility factor is one of the most important factors in choosing the location of a place to live. In addition, locations with a high level of accessibility will have an impact on the development of land use incentives, so locations with a high level of accessibility will have an impact on increasing various new activity centers (Nurbonita & Haryanto, 2017). The 'location theory' generally states that the farther the distance from the city center, the less the value will be. The level of accessibility is influenced by distance, the condition of the transportation infrastructure, and the availability of various means of connecting, including the frequency and the level of security and comfort for going through these routes. The accessibility factor is associated with a trade-off between transportation costs and land prices (Studies, 2019).

#### 3. METHODS

The analytical method used in this study is ensemble learning, which provides the best model for predicting apartment rental prices in Surabaya based on the travelio.com website. Ensemble Learning is a machine learning paradigm in which several models (usually called "weak learners") are trained to solve the same problem and are combined to get better algorithm results.

In this stage, the researcher seeks and learns about the problem to be studied. Then proceed with analyzing the data characteristics of apartments in Surabaya based on the travelio.com website. Next, build a prediction model using the ensemble algorithm to predict apartment rental prices in Surabaya and evaluate the prediction model's performance. Then, proceed with making the results of the website framework for predicting apartment rental prices using Flask in Python based on the prediction results with the level of accuracy that has been obtained (Izquierdo VM 2020). Adjusted to the research objectives, there are two variables used, namely the dependent variable and the independent variable. The dependent variable is a variable whose variation is influenced by the variable. The dependent variable in this study is the apartment rental price per month on a numerical scale. Independent variables are variables whose variations affect other variables. The independent variables in this study are the district, type of unit, bedroom, floor, area, facilities, and accessibility, which consists of the distance of the apartment unit to the nearest airports, station, terminal, campus, and toll gate.

# Jurnal Teknobisnis 2023, Vol. 9(1), 49-59



#### FIGURE 1. Research and Methodology Flow Framework Machine Learning

# 4. RESULTS

#### **Data Collection and Pre-Processing**

Data collection through web scraping using the JSON API with Visual Studio tools, then the data is stored in a Microsoft Excel file with Comma Separated Values (CSV) format. Data from the results of scraping in the form of Excel files obtained as many as 772 apartment units which were still in the form of irregular data, so the researchers carried out cleaning the data with the following steps:

- Checking for missing data and incorrect types for each variable. Such as writing errors using periods, commas, quotes, spaces, and rupiah. Furthermore, data imputation is carried out by looking at the median data;
- Data adjustment by changing the non-numeric/categorical data type to numeric using the label encoder or one hot encoder contained in the sklearn.preprocessing module. This is because the ensemble algorithm can be processed by numeric data. The label encoder process is carried out on the sub-district variable, while the one hot encoder process is carried out on the unit type variable and the completeness of the facilities;
- Checking imbalance data, outlier data, and scaling data using statistical normality tests and standardization so that statistically, it can be processed better in the model;

#### **Prediction of Surabaya City Apartment**

• Training Multiple Models (bagging, boosting, and stacking) followed by model performance evaluation.

|      | Lokasi          | Jenis<br>Unit | Jumlah<br>Bedroom | Luas<br>Unit | Kelengkapan       | Lantai<br>Unit | Kamar<br>Mandi | Harga       | Jarak dari<br>Kampus (km) | Jarak dari<br>Bandara (km) | Jarak dari<br>Stasiun (km) | Jarak dari<br>Terminal (km) | Jarak dari<br>Tol (km) |
|------|-----------------|---------------|-------------------|--------------|-------------------|----------------|----------------|-------------|---------------------------|----------------------------|----------------------------|-----------------------------|------------------------|
| 0    | Mulyorejo       | studioroom    | 1.0               | 24.73        | Full<br>Furnished | 18             | 1.0            | Rp4,196,664 | 3.2                       | 21.5                       | 5.5                        | 15.3                        | 10.2                   |
| 1    | Wiyung          | studioroom    | 1.0               | 21.32        | Full<br>Furnished | 7              | 1.0            | Rp3,210,418 | 4.8                       | 21.4                       | 11.4                       | 6.0                         | 2.3                    |
| 2    | Gunung<br>Anyar | studioroom    | 1.0               | 33.00        | Full<br>Furnished | 4              | 1.0            | Rp4,012,413 | 5.6                       | 8.5                        | 14.9                       | 11.4                        | 12.7                   |
| 3    | Mulyorejo       | 2 BR          | 3.0               | 38.00        | Full<br>Furnished | 26             | 1.0            | Rp5,095,251 | 2.4                       | 22.4                       | 9.2                        | 20.7                        | 12.6                   |
| 4    | Genteng         | studioroom    | 1.0               | 59.00        | Full<br>Furnished | 25             | 1.0            | Rp8,294,000 | 4.4                       | 20.8                       | 4.1                        | 10.8                        | 7.4                    |
| 22.5 |                 |               |                   |              |                   |                |                |             |                           |                            |                            |                             |                        |
| 695  | Dukuh<br>Pakis  | 2 BR          | 2.0               | 36.00        | Full<br>Furnished | 2              | 1.0            | Rp4,138,500 | 3.6                       | 24.3                       | 9.8                        | 14.4                        | 2.5                    |
| 696  | Mulyorejo       | studioroom    | 1.0               | 24.00        | Full<br>Furnished | 28             | 1.0            | Rp3,843,180 | 0.5                       | 21.5                       | 5.5                        | 18.5                        | 10.3                   |
| 697  | Sukolilo        | 2 BR          | 2.0               | 36.00        | Full<br>Furnished | 17             | 1.0            | Rp4,447,300 | 1.2                       | 19.3                       | 6.6                        | 16.4                        | 8.2                    |
| 698  | Mulyorejo       | studioroom    | 1.0               | 24.73        | Full<br>Furnished | 18             | 1.0            | Rp4,196,664 | 0.5                       | 21.5                       | 5.5                        | 18.5                        | 10.3                   |
| 699  | Mulyorejo       | studioroom    | 1.0               | 25.87        | Full<br>Furnished | 14             | 1.0            | Rp4,378,495 | 0.5                       | 21.5                       | 5.5                        | 18.5                        | 10.3                   |

700 rows × 14 columns

#### FIGURE 2. Apartment Unit Data Frames

#### **Ensamble Learning**

Ensemble Learning aims to combine the decisions of several learning algorithms to improve the accuracy of results (especially weak learning algorithms) and create a better model to make predictions. Ensemble Learning is divided into three models: Bagging, Boosting, and Stacking, where machine learning predicts apartment rental prices by writing code using Python language with Jupyter Notebook software.

After cleaning the data, which was initially 772 data, it became 700 by removing some outlier data, then dividing the data into 2, namely training and testing; training data is used to train the algorithm to form a model, and data testing is used to find out the algorithm that has been trained and formed previously on the training data. The division of the training data must be largerthan the testing; if the testing is more significant, there will be overlap, and the model will not "learn." The proportion used this time was 490 training data (70%) and 210 testing data (30%) taken randomly.

#### Bagging

Bagging is an ensemble learning method that only uses one base model type by conducting parallel and independent learning on each base model and then combining themto get the best results.

|    |                                 | Dat          | a Training    | (70%)                   | Data Testing (30%) |               |                         |  |  |
|----|---------------------------------|--------------|---------------|-------------------------|--------------------|---------------|-------------------------|--|--|
| No | Models                          | MAE<br>(IDR) | RMSE<br>(IDR) | R <sup>2</sup><br>Score | MAE (IDR)          | RMSE<br>(IDR) | R <sup>2</sup><br>Score |  |  |
| 1  | Support<br>Vector<br>Regression | 403,589      | 327,93<br>7   | -0.019                  | 433,937            | 369,200       | -0.004                  |  |  |

#### **TABLE 1.** Comparison Data Training and Data Testing of Bagging Model

| 2 | Decision Tree<br>Regression    | 113,070 | 73,084      | 0.919 | 318,798 | 213,042 | 0.457 |
|---|--------------------------------|---------|-------------|-------|---------|---------|-------|
| 3 | Linear<br>Regression           | 330,082 | 263,47<br>9 | 0.318 | 426,874 | 354,523 | 0.027 |
| 4 | Random<br>Forest<br>Regression | 229,183 | 185,696     | 0.671 | 329,879 | 271,710 | 0.419 |

From comparisons by selecting the base model formed and looking at the minor error values seen by MAE and RMSE. The training data shows that the best bagging ensemble learning model is followed by looking at the testing data of the model. Table 1 shows that decision tree regression has a minor error with an MAE value of IDR 318,798 and an RMSE of IDR 213,042 compared to the support vector regression, linear regression, and random forest regression models. The formation of the decision tree regression model obtained an R2 value in the data testing of 45.76%; in this case, the independent variables in the testing data affect the dependent variable in the form of apartment rental prices of 45.76%, while 54.24% are influenced by other factors not included in the model. However, the decision tree regression model is overfitting because the error value in the training data is very low compared to the error value in the data testing, which tends to be high. The error rate is still high with the MAE, RMSE, and R2 values, and analysis with other models is needed.

#### Boosting

Boosting is an ensemble learning method that uses only one base model type bycarrying out sequential adaptive learning (the result of a base model depends on the results of the previous base model) and then combining to get the best results.

|    |  | Data    | Training (70 | %)             | Data    | Data Testing (30%) |                |  |  |
|----|--|---------|--------------|----------------|---------|--------------------|----------------|--|--|
| No | Models                                     | MAE     | RMSE         | R <sup>2</sup> | MAE     | RMSE               | R <sup>2</sup> |  |  |
|    |  | (IDR)   | (IDR)        | Score          | (IDR)   | (IDR)              | Score          |  |  |
| 1  | Boosting- Gradient<br>Boosting Regression  | 113,189 | 74,614       | 0.919          | 299,522 | 203,894            | 0.521          |  |  |
| 2  | Boosting - There is Boosting<br>Regression | 239,748 | 200,318      | 0.640          | 339,188 | 270,489            | 0.386          |  |  |
| 3  | Bagging - Decision Tree<br>Regression      | 149,769 | 106,992      | 0.868          | 248,101 | 181,410            | 0.618          |  |  |
| 4  | Bagging - Random Forest<br>Regression      | 160,772 | 128,612      | 0.838          | 528,473 | 432,362            | -<br>0.490     |  |  |

| TADIE 2  | Comparison | Data Tra | ining and | Data Tacti | na of D | Poorting | Madal |
|----------|------------|----------|-----------|------------|---------|----------|-------|
| IADLE Z. | COMDANSON  | Dala Ha  | ininu anu | Data restr |         | DOSLING  | wouer |
|          |            |          |           |            |         |          |       |

Table 2 shows that the bagging decision tree regression model has the smallest error with an MAE value of IDR 248,101 and an RMSE of IDR 181,410 compared to the gradientboosting regression model, there is boosting regression and bagging-random forest regression. The formation of the bagging decision tree regression model obtained an R2 value in the testing data of 61.85%; in this case, the independent variables in the

#### **Prediction of Surabaya City Apartment**

testing data affect the dependent variable in the form of apartment rental prices by 61.85%, while 38.15% are influenced by other factors not included in the model.

#### Stacking

Stacking is a method that uses several base models by conducting parallel and independent learning on each base model and then combining them using a meta-learning algorithm to provide output from the results of the combined base models.

|    |   | Data         | Training (70  | %)             | Data Testing (30%) |               |                |  |
|----|---|--------------|---------------|----------------|--------------------|---------------|----------------|--|
| No | Models  | MAE<br>(IDR) | RMSE<br>(IDR) | R <sup>2</sup> | MAE<br>(IDR)       | RMSE<br>(IDR) | R <sup>2</sup> |  |
| 1  | Stacking - LR, RF, GBR<br>Regression Meta LR      | 144,106      | 108,761       | 0.843          | 220,791            | 191,776       | 0.621          |  |
| 2  | Stacking - LR, RF, DTR<br>Regression Meta LR      | 160,971      | 123,434       | 0.848          | 247,513            | 180,626       | 0.620          |  |
| 3  | Stacking - GBR, RF, DTR<br>Regression Meta GBR    | 205,845      | 144,340       | 0.752          | 282,700            | 201,965       | 0.504          |  |
| 4  | Stacking - GBR, RF, DTR, LR<br>Regression Meta RF | 204,871      | 145,402       | 0.754          | 268,494            | 201,018       | 0.553          |  |
| 5  | Stacking - GBR, RF, DTR, LR<br>Regression Meta LR | 166,763      | 130,789       | 0.837          | 243,401            | 179,432       | 0.632          |  |

**TABLE 3.** Comparison Data Training and Data Testing of Stacking Model

Based on Table 3 shows that the stacking model with a combination of gradient boosting regression, random forest, decision tree regression, linear regression, meta linear regression has the smallest error with an MAE value of IDR 243,401 and an RMSE of IDR179,432 compared to other stacking models. The formation of a stacking model with a combination of gradient boosting regression, random forest, decision tree regression, linear regression, meta linear regression, linear regression, meta linear regression obtained an R2 value in data testing of 63.28%; in this case, the independent variable in the data testing affects the dependent variable in the form of apartment rental prices of 63.28%, while 36.72% is influenced by other factors not included in the model.

#### Model Performance Evaluation

After analyzing the model, evaluating the model to get the best performance in determininga prediction by comparing the bagging, boosting, and stacking models from the MAE, RMSE, and R<sup>2</sup> values. The smaller the MAE and RMSE values, the smaller the error rate. A comparison of theanalysis results shows that the stacking model with a combination of gradient boosting regression, random forest, decision tree regression, linear regression, and meta-linear regression performs better on training data and data testing. The stacking model has an average MAE and RMSE valuebelow IDR 250,000 and an R2 score above 0.5, so it can be concluded that it has better performance than the Bagging and Boosting Models.

## Variable Importance

Variable Importance used to find out how much the contribution of the sub-district locationvariable, the number of bedrooms, the location of the unit floor, the area of the unit, the type of unit, the bathroom, the completeness of the facilities and the accessibility of the unit from the campus, toll gates, stations, airports, and terminals to predicting apartment rental prices. The greater the variable's value, the greater the importance of this variable in predicting apartment rental prices.



#### FIGURE 3. Variable Importance

Figure 3 shows that the unit area variable gives the most significant role in determining the prediction of apartment rental prices, which is equal to 0.327. Then followed by the completeness of the facility variable by 0.199, the distance from the station to the apartment unit by 0.129, the distance from the toll gate to the unit location by 0.094, and the sub-district location by 0.074.

#### **5. CONCLUSIONS**

In conclusion, this paper addresses that the general description is related. The characteristics of the 772-apartment data in Surabaya City based on the travelio.com website include Apartments in Dukuh Pakis Subdistrict having the highest number, 158 apartment units, and the lowest in Bubutan Subdistrict. Building a prediction model using the ensemble algorithm to predict apartment rental prices in Surabaya using a comparison of 3 (three) models: bagging, boosting, and stacking. Where is the bagging model obtained decision tree regression with the slightest error with an RMSE value of IDR 213,042 and an R2 value of 45.76%. The boosting model found that the decision tree regression bagging model had a minor error with an RMSE value of IDR 181,410 and an R2 value of 61.85%. The model performance evaluation shows that the stacking model with a combination of gradient boosting regression (GBR), random forest (RF), decision tree regression (DTR), linear regression (LR) meta-linear regression (LR) has a minor error with an RMSE value of ADR 179,432 and an R2 value of 63.28%. The stacking model has an average MAE and RMSE value below IDR 250,000 and R2 score above 0.5. According

to that it performs better for predicting apartment rental prices in Surabaya City compared to the single model bagging and boosting models.

# REFERENCES

- Agmelina, E. N., & Ariastita, P. G. (2017). Faktor Pemilihan Lokasi Apartemen Berdasarkan Preferensi Pemerintah di Surabaya Metropolitan Area. *Jurnal Teknik ITS*, 6(2). https://doi.org/10.12962/j23373539.v6i2.24967
- Angelia Gaina, P., & Hariyanto, T. (2014). Evaluasi Perkembangan Dan Persebaran Pembangunan Apartemen Sesuai Dengan RTRW Surabaya.
- Badan Pusat Statistik Kota Surabaya. (2022).
- De Chiara, J., & M. Crosbie, J. (2001). *Time Saver Standards for Residential Development*. Mc Graw Hill Book Companies Inc.
- Kim, K.-O., & Kim, Y.-H. (2015). A Study on Forecasting Model for Apartment Housing Price Index Reflecting Model Uncertainty\*: Focused on BMS, BMA. *Journal of Real Estate Analysis*, 1(1e). https://doi.org/10.30902/jrea.2015.1.1e.27
- Krisnaputri, N. A. (2016). Pola Pemilihan Lokasi Pembangunan Apartemen di Surabaya oleh Pengembang.
- Mulyahati, I. L. (2020). Implementasi Machine Learning Prediksi Harga Sewa Apartemen Menggunakan Alogritma Random Forest Melalui Framework Website Flask Python.
- Nurbonita, R., & Haryanto, R. (2017). Analisis Lokasi Dan Fasilitas Apartemen Kalibata City Serta Implikasinya Terhadap Harga Jual Dan Harga Sewa Yang Ditawarkan. *Teknik Perencanaan Wilayah Kota*, 6(1).
- Peraturan Daerah Kota Surabaya Nomor 12 Tahun 2014 Pasal 16. (2014).
- Purnamasari, A., Herulambang, W., & Adityo, R. D. (2018). Forecasting Sales Prices Appartment Using Fuzzy Tsukamoto (case Study My Tower Apartment). *JEECS (Journal of Electrical Engineering and Computer Sciences)*, *3*(1), 363–368. https://doi.org/10.54732/jeecs.v3i1.141
- Rencana Kerja Pemerintah Daerah Kota Surabaya Tahun 2022. (2022).
- Sri Duto Wisnugroho. (2021). Penerapan Konsep Toleransi Dari Sebuah Keberagaman Untuk Rumah Susun Sederhana Sewa (Rusunawa) Khusus Keluarga, Untuk Pekerja Berpenghasilan Menengah Kebawah Di Surabaya, Jawa Timur. *Sri Duto Wisnugroho*.
- Studies, J. C. for H. (2019). The State of the Nation's Housing 2019. JOINT CENTER FOR HOUSING STUDIES.

# How to cite this article:

Komaruljannah, E., & Fatichah, C. (2023). Prediction of Surabaya City Apartment Rental Price Using the Ensemble Method. *Jurnal Teknobisnis*, *9*(1): 49-59. DOI: 10.12962/j24609463.v9i1.930