

PENERAPAN PEMBELAJARAN MESIN MENGGUNAKAN ALGORITMA REGRESI LINIER UNTUK PREDIKSI HARGA RUMAH

Mina Ismu Rahayu¹, Nurul Hikmah Rafika Dewi², Herna Gunawan³

¹²³STMIK BANDUNG

Sekolah Tinggi Manajemen dan Informatika Bandung
JL. Cikutra No. 113, Bandung 40007, INDONESIA

Contact Address:

¹mina@stmik-bandung.ac.id

ABSTRAK

Harga properti terindikasi tumbuh secara meningkat di pasar primer untuk setiap tahunnya berdasarkan survei Bank Indonesia (BI) untuk Survey Harga Properti Residensial (SHPR) pada kuartal I-2022. Kenaikan IHPR terjadi untuk hampir seluruh tipe rumah. Pertumbuhan untuk Tipe rumah kecil sebesar 2,01%, Pertumbuhan untuk Tipe rumah menengah sebesar 2,18%, dan Pertumbuhan untuk Tipe rumah besar sebesar 1,11%. Sedangkan pada kuartal II-2022, pertumbuhan diperkirakan terbatas sebesar 1,16%. Prediksi harga properti akan sangat dibutuhkan masyarakat terutama masyarakat yang berminat memiliki properti dengan spesifikasi yang diinginkan. Penelitian ini melakukan prediksi harga rumah dengan menerapkan pembelajaran mesin dengan studi kasus data di Pondok Sabilulungan Permai. Data yang digunakan merupakan data tahun 2018-2023. Tujuan penelitian ini adalah membangun aplikasi prediksi harga rumah. Menggunakan RMSE (*Root Mean Squared Error*) untuk penghitung error. Hasil prediksi menggunakan regresi linier memiliki error sebesar 0.66. Berdasarkan hasil model regresi linier didapatkan akurasi sebesar 90.08% yang menunjukkan bahwa sistem ini sudah baik.

Kata Kunci : Regresi Linier, Prediksi Harga Rumah, RMSE, Akurasi, Aplikasi.

ABSTRACT

The price of properties has been indicated to increase annually in the primary market, according to Bank Indonesia's (BI) Residential Property Price Survey (SHPR) for the first quarter of 2022. The IHPR (Residential Property Price Index) rose across almost all house types. Growth for small house types was 2.01% year-on-year, for medium house types 2.18%, and for large house types 1.11%. Meanwhile, in the second quarter of 2022, growth is expected to be limited to 1.16%. Property price predictions will be highly sought after by the public, especially those interested in acquiring properties with their desired specifications. This study predicts house prices by applying machine learning with a case study of data from Pondok Sabilulungan Permai. The data used spans from 2018 to 2023. The objective of this study is to develop a house price prediction application. RMSE (Root Mean Squared Error) was used to calculate errors. The prediction results using linear regression showed an error of 0.66. Based on the linear regression model results, an accuracy of 90.08% was achieved, indicating that the system is performing well.

Keywords: Linear Regression, House Prices, RMSE, Accuracy, Application *Invalid source specified..*

1. Pendahuluan

Rumah merupakan bangunan yang memiliki aspek ekonomi dan dapat dijadikan sebagai investasi jangka panjang, yang merupakan aspek yang harus diperhatikan dalam hidup [1]. Harga properti terindikasi tumbuh secara meningkat di pasar primer untuk setiap tahunnya berdasarkan survei Bank Indonesia (BI) untuk Survey Harga Properti Residensial (SHPR) pada kuartal I-2022. Kenaikan IHPR terjadi untuk hampir seluruh tipe rumah. Pertumbuhan untuk Tipe rumah kecil sebesar 2,01%, Pertumbuhan untuk Tipe rumah menengah sebesar 2,18%, dan Pertumbuhan untuk Tipe rumah besar sebesar 1,11%. Sedangkan pada kuartal II-2022, pertumbuhan diperkirakan terbatas sebesar 1,16%.

Oleh karena itu penerapan sistem prediksi dibutuhkan untuk memudahkan pihak pengembang dalam memprediksi harga rumah. Untuk menetapkan rencana, yang terutama didasarkan pada proyeksi atau prakiraan, prediksi sangatlah penting. Untuk digunakan sebagai dasar pilihan atau untuk membuat aturan yang penting untuk kelangsungan bisnis, organisasi atau badan usaha dapat melakukan prediksi kemungkinan yang terjadi di masa depan [2].

1.1 Rumusan Masalah

1. Bagaimana membuat model dari data yang sudah ada untuk memprediksi harga rumah dengan tingkat akurasi yang tinggi?
2. Bagaimana membangun aplikasi untuk memprediksi harga rumah dan menghasilkan rekomendasi ?

1.2 Tujuan

1. Memperoleh model hasil pembelajaran mesin agar dapat digunakan untuk memprediksi harga rumah.
2. Merancang dan menerapkan pembelajaran mesin untuk prediksi harga properti dengan menggunakan model regresi linier sehingga mendapatkan rekomendasi sesuai dengan prediksi harga rumah.

3. Batasan Masalah

1. Dataset menggunakan data harga rumah di Pondok Sabilulungan Permai dan data harga rumah di Pinus Regency.
2. Penelitian ini hanya berfokus pada prediksi harga rumah.
3. Menggunakan Metode regresi linear berganda.

2. Pembahasan

2.1 Landasan Teori

Tanpa dilatih secara khusus, komputer atau program dapat mempelajari berbagai hal sendiri. Machine learning menurut Putra (2019) adalah teknik penarikan kesimpulan dari data yang menggunakan pendekatan matematis dan berfokus pada pengembangan model matematika yang menangkap pola data [3].

Metode statistik yang disebut analisis regresi menguji korelasi antara dua variabel atau lebih.

Metode statistik untuk mengukur pengaruh berbagai faktor satu sama lain adalah analisis regresi linier. Hal ini memungkinkan untuk memprediksi nilai antar variabel dengan lebih tepat. Dua jenis variabel dibedakan menggunakan analisis regresi [4]:

1. Variabel Y merupakan variabel yang keadaanya dipengaruhi oleh keadaan variabel lain.
2. Variabel X merupakan variabel bebas yang keadaanya tidak dipengaruhi variabel lain.

Rumus Regressi Linier :

$$\hat{Y} = b_0 + b_1X_1 + b_2X_2 + b_3X_3 \quad (1)$$

Keterangan:

\hat{Y} = variabel dependen (variabel tak terikat)

X_1, X_2, X_3 = variabel independen (variabel terikat)

b_0 = konstanta

b_1, b_2, b_3 = koefisien regresi

Istilah “aplikasi” mengacu pada perangkat lunak yang telah diatur untuk menjalankan perintah pengguna agar menghasilkan hasil yang lebih akurat. Frasa ini disesuaikan untuk mencerminkan maksud di balik aplikasi tersebut. Aplikasi biasanya digunakan untuk memproses data sesuai dengan tuntutan komputasi atau pemrosesan data yang diantisipasi untuk memecahkan masalah. Aplikasi biasanya disebut sebagai program komputer dengan fitur unik yang dapat diakses pengguna [5].

2.2 Analisis Teori

Sebelum perancangan sistem dilakukan, penulis harus melakukannya analisis teori dalam penelitian. Yang pertama adalah penulis menganalisis mengenai aspek prediksi harga rumah berdasarkan bukti dari berbagai sumber referensi seperti buku dan internet.

Ada beberapa aspek yang mempengaruhi analisis prediksi harga rumah yaitu :

1. Pengumpulan Data

Pada tahap ini dilakukan pengumpulan dataset yang berasal langsung dari pengembang Pondok Sabilulungan Permai yang berada di Kabupaten Bandung dan juga data pengembang Grand Pinus Regency yang berada di Kota Bandung, adanya data Kota Bandung adalah untuk membandingkan terkait harga dengan karakteristik yaitu Kabupaten dan Kota. Kedua data memiliki field yang sama yaitu Nama Rumah, Tahun, Harga Tanah/m, Luas Bangunan, Luas Tanah, Kamar Mandi,

Kamar Tidur, Suku Bunga dan Harga jual yang kemudian akan di analisis dan dibuat pemodelan melalui jupyter notebook dengan Bahasa phyton.

2. Data Pre-processing

Pada tahap ini dilakukan pre-processing data menggunakan jupyter notebook dengan bahasa kemudian memakai sklearn. Pada tahap ini dilakukan beberapa proses yaitu :

1. Memperbaiki Penulisan

Adanya penulisan yang terdapat kesalahan yaitu pada daftar nama rumah dimana terdapat huruf kecil di awal kata. Dengan menggunakan metode str.title() atau String Title() yaitu untuk mengubah karakter pertama di setiap kata menjadi huruf kapital.

2. Menambah kolom

Menambah kolom akumulasi harga tanah dengan cara :

Harga Tanah/m * Luas Tanah = Harga Tanah

Contoh :

$$2.500.000 * 72 = 180.000.000$$

3. Penerapan Model Regresi Linier

Contoh perhitungan secara manual :

Tabel 3.1 Perhitungan Manual

No	X_1	X_2	X_3	Y
1	1,500,000	21	40	95,220,000
2	1,500,000	25	40	99,820,000
3	1,500,000	30	60	141,105,000
4	1,500,000	36	60	148,005,000
5	2,000,000	45	72	222,318,000
6	2,000,000	50	55	187,795,000
...
300	4,000,000	250	199	1,836,426,000
Σ	843,500,00	32,208	32,190	187,092,362,750

Keterangan :

X_1 = Harga Tanah/m X_2 = Luas Bangunan

X_3 = Luas Tanah Y = Harga Jual

Diketahui :

$n = 300$

$\Sigma X_1 = 843,500,000$

$\Sigma X_2 = 32,208$

$\Sigma X_3 = 32,190$

$\Sigma y = 187,092,362,750$

$\Sigma X_1 X_3 = 92,936,500,000$

$\Sigma X_1 X_3 = 92,366,500,000$

$\Sigma X_2 X_3 = 4,304,298$

$\Sigma X_1^2 = 2,539,250,000,000,000$

$\Sigma X_2^2 = 4,587,492$

$\Sigma X_3^2 = 4,683,666$

$\Sigma X_1 Y = 581,684,640,125,000,000$

$\Sigma X_2 Y = 25,778,072,601,250$

$\Sigma X_3 Y = 26,690,279,136,250$

$$X = \begin{bmatrix} 1 & x_{11} & x_{12} & x_{13} \\ 1 & x_{21} & x_{22} & x_{23} \\ 1 & x_{31} & x_{32} & x_{33} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} & x_{n3} \end{bmatrix}$$

$$X'X =$$

$$\begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ x_{11} & x_{21} & x_{31} & \dots & x_{n1} \\ x_{12} & x_{22} & x_{32} & \dots & x_{n2} \\ x_{13} & x_{23} & x_{33} & \dots & x_{n3} \end{bmatrix} \begin{bmatrix} 1 & x_{11} & x_{12} & x_{13} \\ 1 & x_{21} & x_{22} & x_{23} \\ 1 & x_{31} & x_{32} & x_{33} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} & x_{n3} \end{bmatrix}$$

$$= \begin{bmatrix} n & \Sigma X_1 & \Sigma X_2 & \Sigma X_3 \\ \Sigma X_1 & \Sigma X_1^2 & \Sigma X_1 X_2 & \Sigma X_1 X_3 \\ \Sigma X_2 & \Sigma X_1 X_2 & \Sigma X_2^2 & \Sigma X_2 X_3 \\ \Sigma X_3 & \Sigma X_1 X_3 & \Sigma X_2 X_3 & \Sigma X_3^2 \end{bmatrix}$$

$X'X$

$$= \begin{bmatrix} 300 & 843,500,000 & 32,208 & 32,190 \\ 843,500,000 & 2,539,250,000,000,000 & 92,936,500,000 & 92,366,500,000 \\ 32,208 & 92,936,500,000 & 4,587,492 & 4,304,298 \\ 32,190 & 92,366,500,000 & 4,304,298 & 4,683,666 \end{bmatrix}$$

$(X'X)^{-1}$

$$= \begin{bmatrix} 0.055751077968359 & -0.000000015882342 & -0.000029260950156 & -0.000043060935214 \\ -0.000000015882342 & 0.000000000000006 & -0.00000000012382 & -0.00000000000755 \\ -0.000029260950156 & -0.00000000012382 & 0.00001862055468 & -0.00001265950859 \\ -0.000043060935214 & -0.00000000000755 & -0.00001265950859 & 0.00001687766593 \end{bmatrix}$$

$$X'Y = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ x_{11} & x_{21} & x_{31} & \dots & x_{n1} \\ x_{12} & x_{22} & x_{32} & \dots & x_{n2} \\ x_{13} & x_{23} & x_{33} & \dots & x_{n3} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \Sigma y \\ \Sigma X_1 y \\ \Sigma X_2 y \\ \Sigma X_3 y \end{bmatrix} = \begin{bmatrix} 187,092,362,750 \\ 581,684,640,125,000,000 \\ 25,778,072,601,250 \\ 26,690,279,136,250 \end{bmatrix}$$

$$b = (X'X)^{-1}X'Y$$

$$\begin{bmatrix} 0.055751077968359 & -0.000000015882342 & -0.000029260950156 & -0.000043060935214 \\ -0.000000015882342 & 0.000000000000006 & -0.00000000012382 & -0.00000000000755 \\ -0.000029260950156 & -0.00000000012382 & 0.00001862055468 & -0.00001265950859 \\ -0.000043060935214 & -0.00000000000755 & -0.00001265950859 & 0.00001687766593 \end{bmatrix}$$

$$\begin{bmatrix} 187,092,362,750 \\ 581,684,640,125,000,000 \\ 25,778,072,601,250 \\ 26,690,279,136,250 \end{bmatrix}$$

$$b = \begin{bmatrix} -711512577.17097664 \\ 266.75361894 \\ 1534960.18085720 \\ 3917411.02271122 \end{bmatrix}$$

Sehingga

$$\hat{Y} = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3$$

$$= -711512577.17097664$$

$$+ 266.75361894 X_1$$

$$+ 1534960.18085720 X_2$$

$$+ 3917411.02271122 X_3$$

4. Data Training

Dengan menggunakan fungsi *train* untuk data latih sebesar 80% dari jumlah dataset yaitu 216 baris data.

5. Data Testing

Dengan menggunakan fungsi *test* untuk data latih sebesar 20% dari jumlah dataset yaitu 216 baris data.

6. Prediksi

Dengan Contoh perhitungan :

Diketahui :

$$b_0 = -711512577.17097664 \quad b_2 = 1534960.18085720$$

$$b_1 = 266.75361894$$

$$b_3 = 3917411.02271122$$

Ditanya :

$$X_1 = 2.500.000 \quad X_2 = 52 \quad X_3 = 72$$

Jawaban :

$$= -711512577.17097664 + 266.75361894 \\ * 2.500.000 \\ + 1534960.18085720 * 52 \\ + 3917411.02271122 * 72$$

$$\hat{Y} = 317.242.993.2$$

Maka hasil prediksi untuk rumah dengan spesifikasi harga tanah/m = 2.500.000, luas bangunan = 52 dan luas tanah = 72 yaitu sebesar **Rp. 317.242.993.2**.

2.3 Desain dan Tampilan Antarmuka (User Interface)

1. Halaman Home



2. Halaman Prediction

3. Halaman Hasil Prediksi dan rekomendasi

3. Kesimpulan dan saran

4.1 Kesimpulan

Berdasarkan penelitian yang telah dilakukan mengenai implementasi algoritma regresi linier untuk aplikasi prediksi harga rumah, dapat

disimpulkan bahwa :

1. Menghasilkan model pembelajaran

untuk memprediksi harga rumah dengan menggunakan algoritma regresi linier dengan rata-rata skor RMSE sebesar 0.66 dan skor akurasi sebesar 90.08%. Dengan hasil tersebut dapat diketahui bahwa model yang dihasilkan sudah baik untuk melakukan prediksi harga rumah.

2. Membangun aplikasi prediksi harga rumah dengan menerapkan model regresi linier untuk mengetahui harga prediksi rumah sesuai spesifikasi yang diinginkan kemudian dapat menghasilkan rekomendasi sesuai dengan harga prediksi.

4.2 Saran

Sistem prediksi harga rumah dengan menggunakan algoritma regresi linier telah selesai dibuat, tentunya masih terdapat banyak kekurangan. Berikut saran pengembangan aplikasi mendatang :

1. Untuk meningkatkan kualitas prediksi harga rumah, pengembangan model regresi linier dapat melibatkan penambahan variabel independen yang lebih relevan seperti karakteristik tanah, aksestabilitas dll yang diperkirakan memiliki pengaruh signifikan terhadap harga rumah.
2. Untuk meningkatkan penyempurnaan data kualitas prediksi, pastikan data yang digunakan lebih lengkap dan akurat sehingga mempunyai tingkat akurasi yang tinggi.

DAFTAR PUSTAKA

- [1] Mu'tashim, Muhammad Labib, Sekar Ayu Damayanti, Hanan Nadia Zaki, Toni Muhayat, dan Rio Wirawan, "Analisis Prediksi Harga Rumah Sesuai Spesifikasi Menggunakan Multiple Linear Regression," JURNAL INFORMATIK, 2021.
- [2] Hasmawati, Jumadil Nangi, Mutmainah Muchtar, "Aplikasi Prediksi Penjualan Barang Menggunakan Metode K-Nearest

- Neighbor,” *semanTIK*, vol. 3, pp. 151-160, 2017.
- [3] Diantika, Affif Surya, Yuki Firmanto, “IMPLEMENTASI MACHINE LEARNING PADA APLIKASI PENJUALAN PRODUK DIGITAL (STUDI PADA GRABKIOS),” *Jurnal Ilmiah Mahasiswa FEB*, vol. 9, pp. 1689-1699, 2019.
- [4] Shaputra, Rahmana Dwi, “IMPLEMENTASI REGRESI LINIER UNTUK PREDIKSI PENJUALAN DAN CASHFLOW PADA APLIKASI POINT OF SALES KAFE XYZ,” 2 Agustus 2021. [Online]. Available: <https://dspace.uui.ac.id/bitstream/handle/123456789/36071/17523103%20Rahmana%20Dwi%20Shaputra.pdf?sequence=1>. [Accessed Februari 2023].
- [5] Novendri, Muhammad Saed, Ade Saputra, Chandra Eri Firman, “Aplikasi Inventaris Barang Pada MTS Nurul Islam Dumai Menggunakan PHP dan MySQL,” *Jurnal Manajemen dan Teknologi Informasi*, vol. 2, no. e-ISSN : 2528-1062, p. Mei, 2019).