

Prediksi Prospek Harga *Real Estate* di Masa Pandemi dengan 3 Atribut Berbasis Algoritma *Linear Regression*

Gita Ayu Syafarina¹⁾, Tri Wahyu Qur'ana²⁾, Galih Mahalisa³⁾*

^{1,2,3)} Teknik Informatika, Universitas Islam Kalimantan MAB Banjarmasin,
Jl. Adyaksa No 2, Banjarmasin Utara, Banjarmasin Kota, Kalimantan Selatan 70122
Email : gitaayusyafarina@gmail.com, triwahyuqurana@gmail.com,
*galih.mahalisa@gmail.com

Abstrak

*Investasi menjadi salah satu yang paling banyak diminati sebagian orang bahkan selama pandemi, berbagai jenis investasi mulai dari emas, saham, properti hingga real estate. real estat (real estate) atau lahan yasan adalah sebuah istilah hukum yang mencakup tanah bersama dengan apa pun yang tinggal tetap di atas tanah tersebut, seperti bangunan atau proyek. Usaha real estate adalah jenis kegiatan yang sangat potensial untuk jangka panjang. Oleh karena itulah, banyak orang yang tertarik pada bidang ini termasuk dengan cara berinvestasi. Penelitian ini akan mengolah sebuah dataset dengan menggunakan sebuah metode untuk mendapatkan sebuah prediksi harga real estate dengan melakukan pengujian menggunakan metode *Linear Regression* terhadap 3 atribut yang mempengaruhi harga real estate. nilai akurasi yang didapatkan adalah 67,8%. Maka dari itu untuk harga prospek real estate di masa pandemic tidak mengalami kenaikan atau penurunan dari segi harga jual.*

Kata Kunci : Harga, Linear Regression, Prediksi, Real Estate

Abstract

*Investment is one of the most sought after by some people even during the pandemic, various types of investments range from gold, stocks, and property to real estate. real estate (real estate) or real estate is a legal term that includes land along with anything that remains on the land, such as buildings or projects. The real estate business is a type of activity that has great potential for the long term. Therefore, many people are interested in this field, including investing. This study will process a dataset using a method to get a real estate price prediction by testing using the *Linear Regression* method on 3 features that affect real estate prices. the accuracy value obtained is 67.8%. Therefore, the price of real estate prospects during the pandemic does not increase or decrease in terms of selling prices.*

Keywords: Price, Linear Regression, Prediction, Real Estate

1. PENDAHULUAN

Investasi menjadi salah satu yang paling banyak diminati sebagian orang bahkan selama pandemi, berbagai jenis investasi mulai dari emas, saham, properti hingga real estate [1]. real estat (real estate) atau lahan yasan adalah sebuah istilah hukum yang mencakup tanah bersama dengan apa pun yang tinggal tetap di atas tanah tersebut, seperti bangunan atau proyek [2]. Usaha real estate adalah jenis kegiatan yang sangat potensial untuk jangka panjang. Oleh karena itulah, banyak orang yang tertarik pada bidang ini termasuk dengan cara berinvestasi [3].

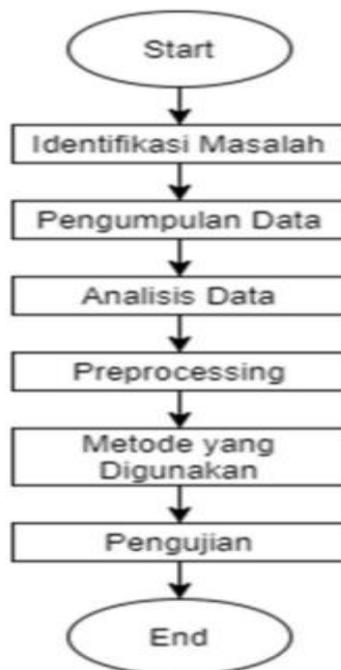
Tinggal di daerah metropolitan adalah janji yang pasti untuk semua karena harga investasi sangat tinggi dan harga tanah, rumah dan fasilitas umum sangat wajar dan harga bisa naik setiap tahun, tapi penelitian ini akan melakukan Analisa jika investasi di masa pandemi khususnya real estate

Referensi yang berkaitan dengan penelitian ini seperti berikut menjelaskan faktor-faktor dari harga rumah, luas bangunan, luas tanah, ruang kamar tidur, garasi, dan kota. Sehingga, hasil penelitian yang didapatkan dengan menggunakan algoritma regresi dengan nilai akurasi sebesar 66% [4]. Penelitian yang lain, menjelaskan bahwa Pengujian keakurasian terhadap hasil prediksi dilakukan menggunakan MSE, RMSE, dan MAPE dapat dikatakan tergolong dalam kategori sangat baik didasari hasil pengujian keakurasian terhadap keseluruhan tipe yang memenuhi standar [5]. Penelitian lain menjelaskan bahwa Pegujian akurasi dilakukan dengan mengukur Root Mean Squared Error (RMSE). Nilai rata-rata akurasi RMSE yang dihasilkan, sebesar 0,432, menunjukkan bahwa variasi nilai yang dihasilkan oleh suatu model prakiraan mendekati akurat [6].

Dalam penelitian ini dilakukan sebuah peramalan atau prediksi yaitu salah satu faktor statistik yang paling penting dalam pengambilan keputusan. Salah satu metode peramalan adalah metode *time series* [7]. Jenis metode yang termasuk dalam metode *time series* adalah *weighted moving average*, *exponential smoothing*, *regresi linier*, dan tren musiman [8]. Keuntungan dari metode regresi linier adalah analisis dapat dilakukan dengan menggunakan beberapa variabel bebas (X), sehingga menghasilkan hasil prediksi yang lebih akurat khususnya di masa pandemi.

2. METODE PENELITIAN

Untuk memprediksi harga real estate adapun pendekatan penelitiannya menggunakan metode kuantitatif, lalu pada tahap metode penelitian yang akan dilakukan adalah dengan pengumpulan data dengan atribut terkait [9], pengolahan data awal (pre-processing data), lalu melakukan penerapan dengan algoritma Linear Regression pada dataset, dan gambaran penelitian yang dilakukan seperti gambar 1 dibawah ini.



Gambar 1 . Tahap Metode Penelitian [9]

2.1 Identifikasi Masalah

Algoritme linear regression merupakan model prediksi atau peramalan yang terdapat lebih dari satu prediktor yang dapat digunakan untuk memprediksi harga jual berdasarkan usia rumah, jarak MRT, dan jumlah toko yang disekitar rumah yang dilakukan pada penelitian ini.

2.2 Pengumpulan Data

Dalam penelitian ini digunakan data yang digunakan dari *Kaggle.com*. Data yang diambil berdasarkan rentang tahun di masa pandemi, mulai tahun 2019 sampai 2020 yang terdiri dari 3 atribut yaitu usia rumah, jarak MRT, dan jumlah toko yang disekitar rumah.

2.3 Analisis Data

Karakteristik dataset terdiri dari 4 variabel. 3 atribut yang dijadikan variabel X atau disebut variabel *dependen* dan 1 variabel yang dijadikan variabel Y atau variabel *independen*, dan data sebanyak 414.

2.4 Preprocessing

Pada tahap *pre-processing*, dilakukan proses *cleaning* data yang digunakan untuk menghilangkan informasi yang tidak diperlukan dalam proses *pre-processing*, yaitu menghilangkan data yang null.

2.5 Metode Linear Regression

Regresi adalah ukuran statistik yang digunakan untuk menentukan kekuatan hubungan antara *variable dependen* (tak bebas) dengan *variabel independen* (bebas) [10]. Metode utama untuk melakukan prediksi yakni membangun model *regresi* dengan mencari hubungan antara satu atau lebih *variable independen* atau prediktor (X) dengan *variable dependen* atau respons (Y) [4]. Linear regression memodelkan hubungan antara variabel skalar dan satu atau lebih variabel penjelas. Dataset penelitian terdiri atas satu variabel *dependen* (Y) dan empat *variable independen* (X) [11]. Variabel *dependen* tersebut adalah atribut Hasil Panen, sedangkan *variable independent* adalah atribut usia rumah, jarak MRT, dan jumlah toko yang disekitar rumah.

2.6. Pengujian

Evaluasi dilakukan menggunakan bantuan perangkat lunak *Jupyter Notebook* menggunakan bahasa *Python* yang digunakan untuk menguji dan melihat tingkat akurasi *algoritme linear regression* terhadap pemrosesan data.

3. HASIL DAN PEMBAHASAN

3.1 Tahap Pengolahan Data

Pengumpulan dataset dengan cara pengunduhan pada API Command kaggle

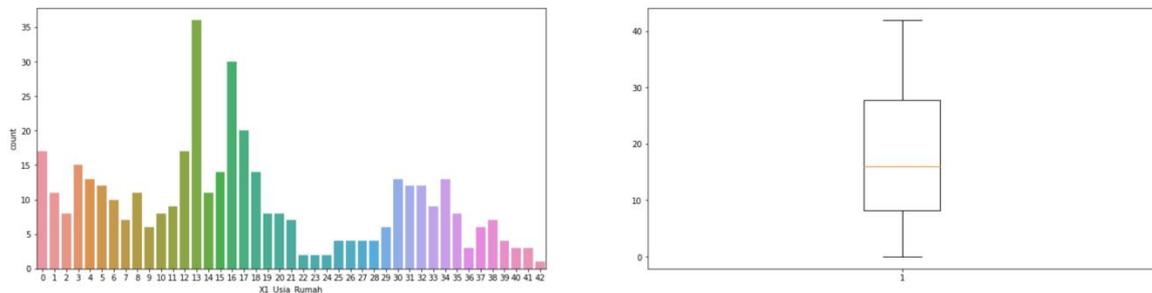
```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 414 entries, 1 to 414
Data columns (total 4 columns):
#   Column          Non-Null Count  Dtype
---  -
0   X1_Usia_Rumah   414 non-null    float64
1   X2_jarak_MRT    414 non-null    float64
2   X3_jumlah_toko  414 non-null    int64
3   Y_harga_unit    414 non-null    float64
dtypes: float64(3), int64(1)
memory usage: 16.2 KB
```

Gambar 2. Dataset

Pengolahan Data dimulai dari identifikasi dataset yaitu memiliki tipe data integer dan data float seperti gambar 2. Setelah melakukan analisa data, tahap selanjutnya melakukan preprocessing, kemudian melakukan split dataset yaitu pembagian data training dan data testing

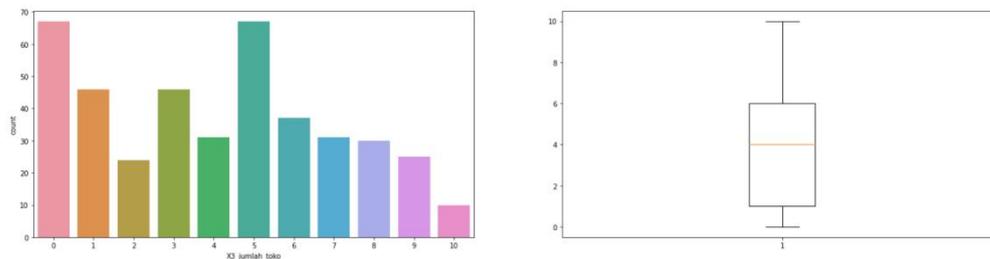
3.2 Melakukan Exploratory Data Analysis (EDA) untuk lebih mengenal data kita dan menemukan insights dari data

Terdapat beberapa atribut yang digunakan yaitu usia rumah, jarak MRT, dan jumlah toko.



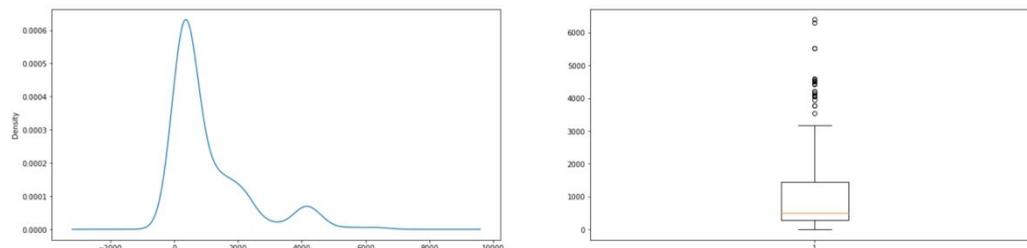
Gambar 3. visualisasi jumlah usia rumah

Pada gambar 3 terlihat bahwa rata rata usia rumah dengan satuan bulan pada real estate adalah paling banyak pada nilai 13



Gambar 4. visualisasi jumlah toko

Pada gambar 4 terlihat bahwa rata rata jumlah toko yang berada disekitar real estate yaitu pada titik 5



Gambar 5. visualisasi jarak MRT

Pada gambar 5 terlihat bahwa rata rata jarak dari MRT menuju real estate adalah 0 atau jarak dekat

3.3. Tahap Pengujian training model menggunakan training data yang sudah di split dan menentukan nilai akurasi

	features	coef_value
0	X1_Usia_Rumah	-0.230110
1	X2_jarak_MRT	-0.005286
2	X3_jumlah_toko	1.346671

```
y_pred = lin_reg.predict(xtest)
lin_reg.score(xtest,ytest)
0.6784662913108155
```

Gambar 6. Hasil Prediksi

Pada gambar 6 terlihat bahwa hasil pengujian dengan Linear Regression menunjukkan nilai prediksi yang baik yaitu sekitar 67,8%.

4. KESIMPULAN

Kesimpulan hasil penelitian dengan menggunakan 3 atribut yang diuji yaitu nilai akurasi yang didapatkan adalah 67,8%. Maka dari itu untuk harga prospek real estate di masa pandemic tidak mengalami kenaikan atau penurunan dari segi harga jual.

5. SARAN

Pada penelitian ini masih memiliki banyak kekurangan sehingga diperlukan rencana pengembangan di penelitian selanjutnya. Beberapa saran yang bisa dilakukan adalah dengan menambahkan lebih banyak atribut atau atribut dari data yang didapatkan

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Lembaga Pusat Penelitian dan Pengabdian Masyarakat (LP2M) Universitas Islam Kalimantan MAB yang telah memberi dukungan materi dan finansial. Serta tidak lupa ucapan terimakasih kepada tim pembantu peneliti yang sudah memberikan waktu dan dukungan untuk ikut serta menyelesaikan penelitian ini.

DAFTAR PUSTAKA

- [1] R. Dwi Rama Jaya, M. Taufiq Ibrahim, R. Riansyah, and W. Nur Happy, "ANALISIS PENGARUH JUMLAH INDUSTRI, JUMLAH PENDUDUK DAN NILAI INVESTASI SEKTOR INDUSTRI KECIL DAN INDUSTRI NON-FORMAL MANUFAKTUR DI KOTA PALEMBANG," *SIBATIK J. J. Ilm. Bid. Sos. Ekon. Budaya, Teknol. dan Pendidik.*, vol. 1, no. 1, pp. 45–52, Dec. 2021, doi: 10.54443/sibatik.v1i1.5.
- [2] W. B. Brueggeman and J. D. Fisher, "Title: Real Estate Finance and Investments," 2011.
- [3] D. L.-J. of P. I. & Finance and undefined 2001, "Real options in real estate development," *emerald.com*.
- [4] M. Labib Mu'tashim, S. A. Damayanti, H. N. Zaki, T. Muhayat, and R. Wirawan, "Analisis Prediksi Harga Rumah Sesuai Spesifikasi Menggunakan Multiple Linear Regression," *ejournal.upnyj.ac.id*, vol. 3, p. 2021.
- [5] G. Najla, A. #1, and D. Fitriana, "Penerapan Metode Regresi Linear Untuk Prediksi Penjualan Properti pada PT XYZ," *J. Telemat.*, vol. 14, no. 2.
- [6] H. W. Herwanto, T. Widiyaningtyas, and P. Indriana, "Penerapan Algoritme Linear Regression untuk Prediksi Hasil Panen Tanaman Padi," 2019.
- [7] N. Nurdiana, S. F. Rodiyansyah, and A. Algifari, "STUDI KOMPARASI ALGORITMA ID3 DAN ALGORITMA NAIVE BAYES UNTUK KLASIFIKASI PENYAKIT DIABETES MELLITUS," *INFOTECH J.*, vol. 6, no. 2, pp. 18–23, Dec. 2020, doi: 10.31949/INFOTECH.V6I2.816.
- [8] W. A. Astuti, "ANALISIS PENGUNGKAPAN CORPORATE SOCIAL RESPONSIBILITY TERHADAP NILAI PERUSAHAAN DI SEKTOR PROPERTY DAN REAL ESTATE," *J. Ris. Akunt.*, vol. 13, no. 2, pp. 186–195, Oct. 2021, doi: 10.34010/jra.v13i2.4779.
- [9] T. S. Sabrila, V. R. Sari, and A. E. Minarno, "Analisis Sentimen Pada Tweet Tentang Penanganan Covid-19 Menggunakan Word Embedding Pada Algoritma Support Vector Machine Dan K-Nearest Neighbor," *Fountain Informatics J.*, vol. 6, no. 2, p. 69, Jul. 2021, doi: 10.21111/fij.v6i2.5536.
- [10] K. Farouq Mauladi, "PERBANDINGAN METODE REGRESI LINEAR DAN NEURAL NETWORK BACKPROPAGATION DALAM PREDIKSI NILAI UJIAN NASIONAL SISWA SMP MENGGUNAKAN SOFTWARE R," 2020.
- [11] T. Indarwati and E. Rimawati, "PENGUNAAN METODE LINEAR REGRESSION UNTUK PREDIKSI PENJUALAN SMARTPHONE," *J. TIKomSiN*, vol. 6, no. 2, 2018, doi: 10.30646/tikomsin.v6i2.369.