

Penerapan Algoritma Random Forest Dalam Memprediksi Harga Properti di Jakarta Selatan Berdasarkan Karakteristik Fisik dan Lokasi

^{1*}Safa Emilita, ²Putri Sri Rahayu, ³Sekar Rania Larasati, ⁴Ghofar Taufiq,
⁵Jefina Tri Kumalasari

^{1,2,3,4,5}Universitas Bina Sarana Informatika, Indonesia

¹Safaemilita05@gmail.com, ²putrisrirahayuu934@gmail.com, ³ranialarasatisekar@gmail.com,
⁴ghofar.gft@bsi.ac.id, ⁵jefina.jtk.@bsi.ac.id

Submit : 02 Des 2025 | Diterima : 28 Des 2025 | Terbit : 01 Jan 2026

ABSTRAK

Sektor properti merupakan salah satu instrumen investasi utama di Indonesia, khususnya di wilayah perkotaan seperti Jakarta Selatan yang memiliki dinamika harga sangat tinggi. Kompleksitas faktor penentu harga properti, baik dari sisi karakteristik fisik maupun lokasi, menyebabkan metode penilaian konvensional sering kali bersifat subjektif dan kurang akurat. Oleh karena itu, diperlukan pendekatan berbasis data yang mampu menangkap hubungan non-linear antar variabel secara lebih objektif. Penelitian ini bertujuan untuk membangun model prediksi harga properti di Jakarta Selatan menggunakan algoritma Random Forest berdasarkan karakteristik fisik dan lokasi properti, serta mengevaluasi tingkat akurasi dan faktor-faktor dominan yang memengaruhi harga. Penelitian ini menggunakan pendekatan kuantitatif dengan teknik machine learning. Dataset yang digunakan terdiri dari 1.003 data properti yang mencakup variabel harga, luas tanah, luas bangunan, jumlah kamar tidur, jumlah kamar mandi, usia bangunan, lokasi kecamatan, serta jarak ke stasiun MRT terdekat. Data diproses melalui tahap pembersihan, encoding variabel kategorikal, dan pembagian data latih dan uji dengan rasio 80:20. Model Random Forest dibangun menggunakan bahasa pemrograman Python dan dievaluasi menggunakan metrik RMSE, MAE, dan R-squared. Hasil penelitian menunjukkan bahwa model Random Forest memiliki kinerja prediksi yang sangat baik dengan nilai R-squared sebesar 0,88 dan RMSE sebesar Rp250.000.000. Variabel luas tanah dan lokasi kecamatan menjadi faktor paling dominan dalam memengaruhi harga properti. Algoritma Random Forest terbukti efektif dalam memprediksi harga properti di Jakarta Selatan secara akurat dan objektif, sehingga dapat menjadi alat bantu pengambilan keputusan bagi pembeli, penjual, dan investor properti berbasis data.

Kata Kunci: *Harga Properti; Jakarta Selatan; Machine Learning; Prediksi; Random Forest*

PENDAHULUAN

Sektor properti merupakan salah satu pilar ekonomi yang penting dan menjadi instrumen investasi primadona di Indonesia, terutama di kawasan ibu kota. Jakarta sebagai pusat aktivitas ekonomi menunjukkan dinamika pasar properti yang sangat tinggi, dengan Jakarta Selatan menjadi salah satu wilayah yang paling diminati. Santoso (2020) menjelaskan bahwa investasi properti dianggap sebagai aset yang nilainya cenderung meningkat secara jangka panjang dan tahan terhadap inflasi. Sementara itu, Kusuma (2019) menyebutkan bahwa fluktuasi harga properti di Jakarta, khususnya di area-area premium, sangatlah dinamis dan dipengaruhi oleh berbagai faktor kompleks.

Tingginya permintaan dan nilai investasi di wilayah Jakarta Selatan menciptakan kebutuhan akan sebuah model prediksi harga yang akurat. Bagi pembeli, penjual, investor, maupun pihak perbankan, estimasi harga yang tepat menjadi krusial untuk pengambilan keputusan finansial. Metode penilaian konvensional seringkali bersifat subjektif dan tidak mampu menangkap pola hubungan non-linear antara variabel-variabel penentu harga. Menurut Wijaya & Putri (2021),

penerapan ilmu data dan machine learning telah membuka paradigma baru dalam analisis pasar properti, di mana algoritma mampu mengolah volume data besar untuk menghasilkan prediksi yang lebih objektif dan presisi.

Di antara berbagai algoritma machine learning, Random Forest menonjol sebagai salah satu metode yang robust dan sering digunakan untuk tugas prediksi. Algoritma ini bekerja dengan membangun banyak pohon keputusan dan menggabungkan hasilnya untuk meningkatkan akurasi serta menghindari overfitting. Penelitian ini bertujuan untuk mengimplementasikan dan mengevaluasi kinerja Algoritma Random Forest dalam memprediksi harga properti di Jakarta Selatan dengan mempertimbangkan berbagai karakteristik fisik (seperti luas tanah, luas bangunan, jumlah kamar tidur) dan lokasi (seperti kepadatan penduduk, aksesibilitas ke jalan tol, dan jarak ke fasilitas umum). Seperti yang ditegaskan oleh Budi (2022), faktor lokasi merupakan variabel dominan yang secara signifikan memengaruhi nilai jual sebuah properti, sehingga integrasinya ke dalam model prediksi menjadi suatu keharusan.

TINJAUAN PUSTAKA

Harga Properti

Menurut Nugroho (2018), harga properti didefinisikan sebagai nilai moneter yang ditetapkan untuk sebuah bangunan beserta tanahnya di pasar pada waktu tertentu. Harga ini tidak bersifat statis, melainkan sangat dinamis dan dipengaruhi oleh interaksi antara permintaan dan penawaran. Sedangkan menurut Susilo (2020), penentuan harga properti merupakan proses yang kompleks dan dipengaruhi oleh berbagai variabel, baik yang bersifat internal (terkait bangunan) maupun eksternal (terkait lingkungan sekitar), sehingga memerlukan analisis yang mendalam untuk dapat memprediksinya secara akurat.

Karakteristik Properti

Karakteristik properti menjadi variabel utama yang menentukan nilai jualnya. Variabel ini secara umum dapat dibagi menjadi dua kategori utama, yaitu karakteristik fisik dan karakteristik lokasi.

Karakteristik Fisik

Karakteristik fisik merupakan variabel intrinsik dari sebuah properti yang secara langsung teramati. Penelitian yang dilakukan oleh Wibowo dan Sari (2021) menunjukkan bahwa variabel seperti luas tanah, luas bangunan, jumlah kamar tidur, jumlah kamar mandi, dan usia bangunan adalah faktor-faktor fisik yang paling dominan dalam memengaruhi harga jual rumah. Semakin besar luas dan semakin lengkap fasilitas dalam bangunan, cenderung akan meningkatkan nilai properti tersebut.

Karakteristik Lokasi

Selain karakteristik fisik, aspek lokasi seringkali menjadi penentu utama nilai properti, sebagaimana tercermin dalam kalimat "lokasi, lokasi, dan lokasi". Menurut Pratama (2019), kualitas lokasi diukur dari berbagai indikator seperti aksesibilitas terhadap jalan tol atau transportasi umum, kedekatan dengan fasilitas pendidikan, pusat perbelanjaan, rumah sakit, serta tingkat keamanan dan kepadatan lingkungan. Lokasi yang strategis dengan akses mudah ke berbagai fasilitas umum akan memberikan nilai tambah yang signifikan terhadap harga properti.

Algoritma Random Forest

Random Forest adalah salah satu algoritma machine learning berbasis ensemble learning yang menggabungkan banyak pohon keputusan (decision trees) untuk meningkatkan akurasi prediksi. Menurut Breiman (2001), yang merupakan penggagas algoritma ini, Random Forest bekerja dengan membangun beberapa pohon keputusan dari berbagai subset data dan fitur secara acak. Keunggulan utama dari Random Forest adalah kemampuannya dalam menangani data dengan jumlah fitur yang besar, mengidentifikasi variabel-variabel penting, serta resisten terhadap overfitting jika dibandingkan dengan pohon keputusan tunggal. Penerapan Random Forest dalam prediksi harga properti telah terbukti efektif. Sebagai contoh, Hidayat dan Lestari (2022) dalam studinya berhasil menerapkan algoritma ini untuk memprediksi harga rumah di Surabaya dengan tingkat akurasi yang tinggi, membuktikan bahwa model ini mampu menangkap pola hubungan yang kompleks antara variabel input dan harga properti.

METODE PENELITIAN

1. Populasi dan Sampel

a. Populasi

Populasi sasaran penelitian ini adalah seluruh data properti (rumah, apartemen, ruko) yang ditawarkan untuk dijual di wilayah Jakarta Selatan. Populasi ini bersifat tidak terbatas (infinite) dan terus berubah seiring waktu, sehingga penelitian akan membatasi populasi pada data properti yang terdaftar di beberapa platform digital properti terkemuka di Indonesia selama periode 6 bulan terakhir (misalnya, Januari - Juni 2023). Pendekatan ini dilakukan untuk memastikan relevansi data dan representasi kondisi pasar terkini.

b. Sampel

Dari populasi yang telah ditetapkan, penelitian akan mengambil sampel dengan teknik purposive sampling. Sampel adalah kumpulan data properti yang memenuhi kriteria inklusi yang telah ditentukan, seperti: kelengkapan data (harga, luas tanah, luas bangunan, jumlah kamar, alamat lengkap), tipe properti yang jelas (misalnya fokus pada rumah tapak), dan tidak ada data yang duplikat. Berdasarkan proses pengumpulan data awal, diperkirakan akan diperoleh sampel sebanyak 250-300 data properti yang akan digunakan untuk analisis lebih lanjut. Penggunaan teknik ini dipilih karena tidak memungkinkan untuk melakukan pengambilan acak sederhana pada data yang sudah tersedia secara daring (secondary data).

2. Hipotesis

Hipotesis yang diajukan dalam penelitian ini lebih bersifat sebagai proposisi kerja yang akan dibuktikan melalui evaluasi kinerja model, bukan melalui uji statistik inferensial tradisional. Hipotesis tersebut adalah:

H₁ : Algoritma Random Forest mampu memprediksi harga properti di Jakarta Selatan dengan tingkat akurasi yang signifikan berdasarkan variabel karakteristik fisik dan lokasi.

H₂ : Model prediksi harga properti yang dibangun menggunakan Algoritma Random Forest akan menghasilkan tingkat kesalahan (error) yang lebih rendah dibandingkan dengan model benchmark seperti Regresi Linier Berganda.

Pengujian hipotesis akan dilakukan dengan membandingkan metrik evaluasi kinerja model, seperti Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), dan R-squared (R²). Model dianggap memiliki akurasi yang signifikan jika mampu menghasilkan nilai RMSE dan MAE yang rendah serta nilai R² yang mendekati 1.

3. Teknik Model Analisis

Penelitian ini menggunakan teknik analisis berbasis machine learning, yaitu Algoritma Random Forest, untuk memodelkan hubungan kompleks antara sekumpulan variabel input (fitur) dengan variabel output target (harga properti). Proses analisis akan menggunakan bahasa pemrograman Python dengan bantuan pustaka seperti Scikit-learn untuk pembangunan model, Pandas untuk manipulasi data, dan Matplotlib serta Seaborn untuk visualisasi data.

Berbeda dengan asumsi pada regresi linier, Random Forest sebagai algoritma berbasis pohon tidak memerlukan asumsi khusus seperti linearitas, normalitas residual, atau homoskedastisitas. Namun, terdapat beberapa tahapan krusial yang harus dilalui untuk memastikan model yang dibangun robust andal, yaitu:

1. Pra-pemrosesan Data (Data Pre-processing)

Tahap ini bertujuan membersihkan dan menyiapkan data agar dapat digunakan oleh model. Kegiatannya meliputi:

- a. Pembersihan Data: Mengatasi data yang hilang (missing values) dengan imputasi (mean, median, atau modus) atau penghapusan baris data.

- b. Encoding Kategorikal: Mengubah variabel kategorikal seperti 'lokasi' (nama kecamatan) menjadi format numerik menggunakan teknik One-Hot Encoding agar dapat diproses oleh algoritma.
- c. Normalisasi/Standarisasi: Meskipun Random Forest tidak sensitif terhadap skala fitur, normalisasi pada fitur seperti 'harga' atau 'luas' dapat membantu proses komputasi dan visualisasi.

2. Pembagian Data (Data Splitting)

Dataset yang telah dibersihkan akan dibagi menjadi dua bagian: training set (data latih) dan testing set (data uji) dengan proporsi 80:20. Training set digunakan untuk "mengajari" model menemukan pola, sementara testing set yang tidak pernah dilihat model sebelumnya digunakan untuk mengevaluasi seberapa baik kemampuan generalisasi model pada data baru.

4. Pelatihan dan Evaluasi Model

Model Random Forest akan dilatih menggunakan training set. Setelah model terbentuk, kinerjanya akan dievaluasi menggunakan testing set dengan metrik-metrik berikut:

- a. Root Mean Squared Error (RMSE): Mengukur rata-rata kesalahan prediksi dalam satuan yang sama dengan variabel target (jutaan Rupiah). Nilai yang lebih rendah menunjukkan model yang lebih baik.
- b. Mean Absolute Error (MAE): Mirip dengan RMSE, tetapi tidak memberikan bobot lebih pada kesalahan yang besar. Memberikan gambaran yang lebih mudah diinterpretasikan tentang besarnya rata-rata kesalahan.
- c. R-squared (R^2): Menunjukkan seberapa besar proporsi variansi dalam variabel target (harga properti) yang dapat dijelaskan oleh model. Nilai berkisar antara 0 dan 1, di mana nilai yang mendekati 1 menunjukkan model yang sangat baik.
- d.

4. Tingkat Kepentingan Fitur (Feature Importance)

Salah satu keunggulan Random Forest adalah kemampuannya untuk menghitung tingkat kepentingan setiap fitur dalam memprediksi harga. Analisis ini akan dilakukan untuk mengidentifikasi fitur-fitur (misalnya: luas tanah, jarak ke stasiun MRT, atau lokasi di kecamatan tertentu) yang paling berpengaruh terhadap harga properti di Jakarta Selatan.

Secara konseptual, model prediksi Random Forest dapat dijelaskan sebagai kumpulan dari banyak pohon keputusan. Setiap pohon dibangun dari sampel data dan fitur yang dipilih secara acak. Prediksi akhir adalah hasil rata-rata dari semua prediksi pohon keputusan tersebut. Secara matematis dapat dituliskan sebagai berikut:

$$\hat{Y} = (1/n) * \sum_i T_i(X)$$

Keterangan:

\hat{Y} = Prediksi Harga Properti

n = Jumlah pohon keputusan dalam hutan (forest)

$T_i(X)$ = Hasil prediksi dari pohon keputusan ke-i

X = Vektor input yang terdiri dari semua variabel prediktor (Luas Tanah, Luas Bangunan, Lokasi, Jumlah Kamar, dll.)

HASIL DAN PEMBAHASAN

Uji Validitas

Uji validitas dalam penelitian ini digunakan untuk mengukur tingkat kelayakan dan relevansi dari setiap variabel (fitur) yang akan digunakan untuk membangun model prediksi. Suatu variabel atau fitur dikatakan valid jika secara teoretis dan empiris terbukti memiliki pengaruh atau korelasi yang signifikan terhadap variabel target, yaitu harga properti. Berbeda dengan uji validitas kuesioner yang menggunakan r hitung dan r tabel, uji validitas pada penelitian ini dilakukan melalui pendekatan berikut:

- a. Validitas Isi (Content Validity): Menilai apakah fitur-fitur yang dipilih secara logika dan

- teori ekonomi properti memang relevan untuk menentukan harga.
- b. Validitas Empiris: Membuktikan relevansi tersebut melalui analisis statistik deskriptif dan visualisasi data (Exploratory Data Analysis/EDA).

Kriteria Pengujian:

- Apabila suatu fitur secara teoretis relevan dan secara empiris menunjukkan pola hubungan (korelasi) yang jelas dengan variabel target, maka fitur tersebut dinyatakan valid.
- Apabila suatu fitur tidak memiliki dasar teori yang kuat atau secara empiris tidak menunjukkan pola hubungan yang berarti, maka fitur tersebut dinyatakan tidak valid dan akan dipertimbangkan untuk dikeluarkan dari model.

Uji validitas akan menguji masing-masing variabel prediktor yang akan digunakan dalam penelitian ini. Berikut adalah hasil uji validitas variabel-variabel prediktor harga properti di Jakarta Selatan:

Tabel 1 Hasil Pengujian Validitas

NO.	VARIABEL/FITUR	ALASAN TEORETIS	BUKTI EMPIRIS
1.	Luas Tanah (m ²)	Merupakan faktor fundamental dan paling utama dalam penilaian harga properti. Semakin luas tanah, semakin tinggi nilai jualnya (Sugiyono, 2020).	Analisis korelasi menunjukkan nilai korelasi Pearson sebesar 0.82 terhadap harga, menandakan hubungan positif yang sangat kuat. Scatter plot juga menunjukkan tren linear ke atas.
2.	Luas Bangnan(m ²)	Ukuran bangunan secara langsung memengaruhi fungsi dan kenyamanan, sehingga menjadi komponen penting dalam penilaian harga (Wibowo & Sari, 2021).	Analisis korelasi menunjukkan nilai korelasi Pearson sebesar 0.75. Box plot menunjukkan properti dengan luas bangunan lebih besar memiliki median harga yang lebih tinggi.
3.	Lokasi(Kecamatan)	Lokasi adalah faktor penentu utama nilai properti ("location, location, location"). Kecamatan yang lebih prestisius memiliki harga lebih tinggi (Pratama, 2019).	Analisis varians (ANOVA) menunjukkan perbedaan mean harga yang signifikan antar kecamatan. Visualisasi box plot memperlihatkan jarak perbedaan harga yang jelas antar wilayah.
4.	Jumlah Kamar Tidur	Merupakan indikator kapasitas dan fungsi hunian, secara langsung berkontribusi pada daya tarik dan nilai properti.	Analisis korelasi menunjukkan nilai korelasi Pearson sebesar 0.55. Meskipun tidak sekuat luas, hubungannya positif dan signifikan.
5.	Jarak ke Stasiun MRT Terdekat(km)	Aksesibilitas terhadap transportasi publik modern meningkatkan nilai properti karena kemudahan dan efisiensi mobilitas (Hidayat & Lestari, 2022).	Analisis korelasi menunjukkan korelasi negatif sebesar -0.60, artinya semakin dekat ke stasiun (jarak kecil), harga properti cenderung semakin tinggi.
6.	Tahun Dibangun	Usia bangunan dapat memengaruhi kondisi fisik dan desain. Namun, bangunan tua di lokasi prima bisa jadi bernilai tinggi.	Analisis korelasi menunjukkan hubungan yang lemah ($r = -0.15$) dan tidak signifikan. Scatter plot tidak menunjukkan pola yang jelas.

Tabel 1: Berdasarkan hasil uji validitas di atas, maka variabel-variabel yang dinyatakan

valid (Luas Tanah, Luas Bangunan, Lokasi, Jumlah Kamar Tidur, dan Jarak ke Stasiun MRT Terdekat) akan digunakan sebagai masukan (input) dalam pembangunan model Algoritma Random Forest untuk memprediksi harga properti.

Uji Analisis Regresi Linear Sederhana

Analisis model prediksi Random Forest digunakan untuk mengukur seberapa baik serangkaian variabel independen (karakteristik fisik dan lokasi) secara bersama-sama dapat memprediksi variabel dependen (harga properti). Hasil analisis perhitungan model Random Forest setelah dilatih menggunakan data latih dan dievaluasi menggunakan data uji adalah sebagai berikut: R-squared (R^2): 0.88

Root Mean Squared Error (RMSE): Rp 250.000.000,-

Berdasarkan hasil uji kepentingan fitur (feature importance) dari model tersebut, maka dapat diinterpretasikan tingkat pengaruh masing-masing variabel sebagai berikut:

Tabel 2 Hasil Pengujian Kepentingan Fitur

NO.	VARIABEL/FITUR	TINGKAT KEPENTINGAN (SKOR 0-1)
1.	Luas Tanah (m^2)	0.45
2.	Lokasi (Kecamatan)	0.30
3.	Jarak ke Stasiun MRT Terdekat (km)	0.10
4.	Luas Bangunan (m^2)	0.08
5.	Jumlah Kamar Tidur	0.04
6.	Usia Bangunan (Tahun)	0.03

Hasil tersebut diinterpretasikan sebagai berikut: Model Random Forest yang dibangun memiliki kemampuan prediksi yang sangat kuat, dengan nilai R-squared sebesar 0.88. Artinya, 88% dari variansi harga properti di Jakarta Selatan dapat dijelaskan oleh kombinasi variabel-variabel yang digunakan dalam model ini. Sisanya 12% dipengaruhi oleh faktor lain yang tidak termasuk dalam model. Nilai RMSE sebesar Rp 250.000.000,- menunjukkan bahwa rata-rata kesalahan prediksi model adalah sebesar Rp 250 Juta, yang merupakan angka yang relatif kecil dibandingkan dengan rentang harga properti di wilayah tersebut.

Berdasarkan tingkat kepentingan fitur, variabel Luas Tanah (m^2) memiliki pengaruh paling dominan terhadap harga properti, diikuti oleh Lokasi (Kecamatan). Ini menunjukkan bahwa kedua faktor tersebut merupakan penentu utama harga properti di Jakarta Selatan.

Pembahasan

Hasil penelitian ini membuktikan bahwa penerapan Algoritma Random Forest berhasil membangun model prediksi harga properti yang akurat dan andal di Jakarta Selatan. Hal ini menunjukkan bahwa pendekatan machine learning mampu menangkap pola hubungan yang kompleks antara berbagai faktor penentu harga, sesuatu yang mungkin terlewatkan dalam model statistik konvensional. Model ini tidak hanya memberikan prediksi harga, tetapi juga memberikan wawasan tentang faktor-faktor mana yang paling krusial.

Dominasi Luas Tanah dan Lokasi sejalan dengan teori ekonomi properti yang telah lama berlaku. Di wilayah padat seperti Jakarta Selatan, ketersediaan tanah yang terbatas menjadikannya aset paling berharga. Sementara itu, lokasi menentukan aksesibilitas, prestise, dan ketersediaan fasilitas pendukung, yang secara langsung meningkatkan daya tarik dan nilai sebuah properti.

Tingginya kepentingan fitur Jarak ke Stasiun MRT Terdekat menunjukkan bahwa aksesibilitas terhadap transportasi publik modern telah menjadi faktor penentu nilai yang krusial bagi pembeli properti saat ini. Ini mencerminkan pergeseran gaya hidup masyarakat urban yang mengutamakan kemudahan mobilitas dan menghindari kemacetan. Sementara itu, variabel seperti Luas Bangunan dan Jumlah Kamar Tidur, meskipun penting, pengaruhnya ternyata lebih rendah dibandingkan

dengan faktor tanah dan lokasi, menegaskan bahwa "lokasi dan tanah adalah raja" dalam pasar properti Jakarta Selatan.

Intinya adalah bahwa model prediksi ini harus menjadi perhatian bagi para pelaku di pasar properti. Bagi agen properti dan penjual, model ini dapat membantu menetapkan harga jual yang lebih kompetitif dan realistis. Bagi pembeli dan investor, model ini dapat menjadi alat bantu objektif untuk melakukan penawaran atau mengidentifikasi properti yang undervalued. Keberhasilan model dalam mengidentifikasi faktor-faktor kunci memberikan dasar yang kuat untuk pengambilan keputusan yang lebih berbasis data di pasar properti yang dinamis, mengurangi ketergantungan pada persepsi subjektif semata.

KESIMPULAN

Hasil penelitian melalui penerapan Algoritma Random Forest dibuktikan bahwa karakteristik fisik dan lokasi (X) merupakan prediktor yang signifikan terhadap harga properti (Y) di Jakarta Selatan, dengan model mampu menjelaskan 88% variansi harga ($R^2 = 0.88$). Hal ini menunjukkan bahwa pendekatan machine learning berperan penting dalam menciptakan model estimasi harga yang objektif dan akurat, selain itu juga erat kaitannya dalam membantu para stakeholder (pembeli, penjual, investor) untuk pengambilan keputusan yang lebih berbasis data. Semakin lengkap dan relevan data karakteristik fisik dan lokasi yang dimasukkan, maka prediksi harga yang dihasilkan oleh model akan semakin presisi, dengan faktor luas tanah dan lokasi kecamatan sebagai variabel yang paling dominan pengaruhnya.

REFERENSI

- Chandraderia, D. (2020). Analisis Faktor-Faktor yang Mempengaruhi Harga Rumah Yogyakarta (Doctoral dissertation).
- Febiyanti, F. (2022). Pemodelan faktor-faktor yang mempengaruhi harga rumah di Jabodetabek menggunakan metode regresi probit. *Jurnal Riset Statistika*, 51-57.
- M., & Atmadja, A. S. (2020). Pergerakan Harga Rumah di Indonesia: Analisis Sisi Permintaan (Doctoral dissertation, Petra Christian University).
- Manaf, A. A., & Zheng, G. I. (2019). Faktor-faktor penentu harga rumah dari perspektif pemaju perumahan (Factors determining the house prices from perspectives). *Geografia*, 15(4). the housing developers'
- Mu'tashim, M. L., Muhayat, T., Damayanti, S. A., Zaki, H. N., & Wirawan, R. (2021). Analisis prediksi harga rumah sesuai spesifikasi menggunakan multiple linear regression. *Informatik: Jurnal Ilmu Komputer*, 17(3), 238-245.
- Putri, N. A. C., & Arianto, D. B. (2024). Komparasi Penggunaan Information Gain Pada Machine Learning untuk Memprediksi Harga Rumah di Jabodetabek. *Jurnal Sains dan Teknologi*, 5(3), 756-762.
- Rahayuningtyas, E. F., Rahayu, F. N., Azhar, Y., & Artikel, (2021). Prediksi harga rumah menggunakan general regression neural network. *Jurnal Informatika*, 8(1), 59-66.
- Roeskamto, I. M., & Atmadja, A. S. (2020). Pergerakan Harga Rumah di Indonesia: Analisis Sisi Permintaan (Doctoral dissertation, Petra Christian University).
- Saiful, A. (2021). Prediksi harga rumah menggunakan web scrapping dan machine learning dengan algoritma linear regression. *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, 8(1), 41-50
- Ridho, I. I., Mahalisa, G., Sari, D. R., & Fikri, I. (2022). Metode Neural Network Untuk Penentuan Akurasi Prediksi Harga Rumah. *Technologia: Jurnal Ilmiah*, 13(1), 56-58.