

Model Prediksi Nilai Jual Apartemen Berbasis *Gross Income Multiplier* dan *Machine Learning*

Njo Anastasia^{a,1*}, Wirawan Sunyoto^{b,2}

^{a,b}Universitas Kristen Petra, Surabaya, Indonesia

Email : ^{1*}anas@petra.ac.id, ²B21240011@john.petra.ac.id

*Corresponding Author

Abstrak

Perkembangan teknologi *big data* telah mendorong peningkatan akurasi dalam penilaian properti melalui pendekatan *Machine Learning* (ML). Tujuan penelitian ini mengevaluasi kinerja model ML dalam memprediksi harga jual apartemen di Kota Surabaya. Selain itu, penelitian ini menggunakan *Gross Income Multiplier* (GIM) sebagai indikator investasi properti untuk menilai konsistensi hasil prediksi harga jual yang dihasilkan oleh model terbaik. Data dikumpulkan melalui teknik *web scraping* dari situs *listing* apartemen di Surabaya dalam rentang waktu 19 September - 21 Oktober 2024. Dua model ML, yaitu *Linear Regression* (LR) dan *Artificial Neural Network* (ANN), dikembangkan menggunakan IBM SPSS Modeler dan dievaluasi menggunakan *Mean Absolute Error* (MAE) dan *Root Mean Squared Error* (RMSE). Hasil penelitian menunjukkan bahwa ANN memiliki kinerja prediksi lebih baik dibandingkan LR, dengan variabel "luas bangunan" sebagai prediktor paling signifikan terhadap harga jual apartemen. Selanjutnya, ANN dianalisis menggunakan GIM sebagai indikator berbasis pendapatan sewa untuk mengevaluasi kewajaran nilai investasi. Hasil perbandingan menunjukkan bahwa median harga jual hasil prediksi ANN relatif konsisten dengan nilai GIM historis. Temuan ini menunjukkan ANN tidak hanya efektif dalam prediksi harga jual, tetapi juga berpotensi untuk analisis kelayakan investasi properti. Integrasi model ML dalam sistem penilaian properti diharapkan dapat meningkatkan efisiensi pengambilan keputusan bagi investor dan pengembang serta mendukung perencanaan kota berbasis data di sektor perumahan vertikal.

©2025 the author(s)

This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Riwayat Artikel

Diterima 2025-11-05

Revisi 2026-01-22

Disetujui 2026-05-13

Kata Kunci

Artificial neural network

Gross income multiplier

Machine learning

Penilaian apartemen

Web scraping

Klasifikasi JEL*:

C45, C53, R31, G17

Pendahuluan

Kota Surabaya merupakan salah satu kota metropolitan terbesar di Indonesia yang mengalami pertumbuhan ekonomi dan urbanisasi pesat signifikan. Menurut Darmawan (2024), jumlah penduduk Surabaya pada 2024 tercatat mencapai 3,02 juta jiwa, meningkat dibandingkan tahun

2023. Dalam lima tahun terakhir, rata-rata pertumbuhan tahunan (*Compounded Annual Growth Rate* - CAGR) penduduk sebesar 0,82% menjadikan Surabaya sebagai kota dengan populasi tertinggi di Provinsi Jawa Timur. Pertumbuhan ini membawa tantangan tersendiri, salah satunya terkait ketersediaan dan keterjangkauan hunian di tengah tekanan ekonomi masyarakat. Urbanisasi yang cepat turut menciptakan permasalahan ketenagakerjaan, terutama di kalangan pemuda. Pada Agustus 2020, tingkat pengangguran pemuda di Indonesia mencapai 15,86%, jauh lebih tinggi dibandingkan orang dewasa sebesar 3,81%. Tingginya pengangguran berdampak pada lemahnya perencanaan keuangan jangka panjang, termasuk kemampuan untuk memiliki rumah (Yanindah, 2021). Tekanan inflasi dan stagnasi upah semakin memperarah kondisi tersebut. Meskipun terjadi pertumbuhan ekonomi, upah yang stagnan tidak sebanding dengan kenaikan harga rumah baik untuk disewa maupun disewa (Snapcart, 2025). Akibatnya, banyak keluarga terpaksa menunda pembelian rumah dan mulai mempertimbangkan apartemen sebagai solusi hunian yang lebih terjangkau dan realistis (Bedi & Rahadi, 2024). Kondisi demografis dan tekanan ekonomi tersebut tidak hanya memengaruhi permintaan hunian, tetapi juga meningkatkan ketidakpastian dalam penentuan nilai pasar apartemen, khususnya dalam konteks penilaian aset yang membutuhkan estimasi nilai yang objektif dan dapat dipertanggung jawabkan.

Sejalan dengan pertumbuhan preferensi hunian ini, permintaan terhadap apartemen terus mengalami peningkatan. Colliers (2024) mencatat adanya pertumbuhan permintaan sebesar 0,5% pada Tahun 2024, dan diperkirakan akan meningkat hingga 87% pada tahun 2026. Salah satu pendorongnya adalah kebijakan insentif PPN yang diterapkan pemerintah sejak 2021. Insentif sebesar 100% kembali diberlakukan pada September - Desember 2024, dan berhasil menarik minat pembeli, termasuk mereka yang sebelumnya belum berencana untuk membeli properti. Minat terhadap apartemen juga dipengaruhi oleh berbagai pertimbangan seperti kondisi fisik, harga, dan lokasi (Lauren, 2023). Fenomena serupa terjadi di Amerika Serikat, di mana keterbatasan lahan dan mahalnya hunian di sekitar kampus mendorong mahasiswa untuk menyewa apartemen sebagai pilihan yang lebih praktis dan fungsional (Ong et al., 2013). Hal ini menunjukkan bahwa tren beralih ke hunian vertikal tidak hanya dipicu oleh kondisi lokal, tetapi juga merupakan respons global terhadap tantangan urbanisasi dan ekonomi. Peningkatan aktivitas pasar apartemen ini menuntut adanya pendekatan penilaian yang mampu menangkap dinamika harga secara akurat, terutama ketika data transaksi aktual terbatas dan pasar berada dalam kondisi yang fluktuatif.

Dalam praktik penilaian apartemen, kondisi pasar yang dinamis sering kali menyulitkan penilai dalam menentukan nilai pasar yang representatif. Keterbatasan data transaksi pembanding, heterogenitas karakteristik unit, serta potensi bias dalam penyesuaian harga menyebabkan pendekatan penilaian konvensional menghadapi tantangan dalam menghasilkan estimasi nilai yang konsisten dan objektif. Kondisi ini menjadi semakin relevan di pasar apartemen Surabaya, di mana transaksi tidak selalu terjadi secara merata antar wilayah dan segmen harga. Berbagai penelitian terdahulu menunjukkan bahwa harga jual apartemen dipengaruhi oleh kombinasi faktor struktural, lokasi, lingkungan, serta fasilitas dan manajemen. Dari sisi karakteristik struktural, variabel seperti luas unit, jumlah kamar, kualitas finishing, dan posisi lantai terbukti signifikan meningkatkan harga jual apartemen (Hoxha et al., 2021; Museleku, 2022). Studi di Nairobi misalnya menemukan bahwa ukuran apartemen dan kualitas interior merupakan determinan utama nilai apartemen (Museleku, 2022), sementara penelitian di Prishtina menegaskan pentingnya luas unit dan aksesibilitas dalam menentukan harga (Hoxha et al., 2021). Selain itu, lokasi juga menjadi faktor dominan, di mana akses menuju pusat bisnis, jaringan transportasi publik, serta kedekatan dengan fasilitas komersial terbukti meningkatkan nilai apartemen (de la Torre & Chen, 2024).

Lebih lanjut, variabel lingkungan dan amenitas juga menunjukkan pengaruh signifikan. Penelitian di Seoul menemukan bahwa keberadaan pemandangan alam (*natural view*) dapat menaikkan harga apartemen, meskipun dampaknya tidak seragam di semua segmen pasar (Lee et al., 2020). Hasil serupa ditunjukkan oleh Ke et al. (2022) yang mengungkapkan bahwa lanskap

ekologis dan kualitas lingkungan sekitar berkontribusi terhadap peningkatan harga hunian. Selain itu, faktor manajemen dan reputasi pengembang turut berperan penting dalam segmen apartemen kelas atas, sebagaimana ditunjukkan oleh studi di Hanoi yang menegaskan bahwa kualitas layanan dan kompetensi pengembang secara positif memengaruhi harga jual apartemen. Dengan demikian, meskipun terdapat variasi konteks antar negara, pola yang konsisten dari temuan tersebut menunjukkan bahwa harga apartemen ditentukan oleh kombinasi variabel struktural, lokasi, lingkungan, serta fasilitas dan manajemen. Meskipun berbagai studi tersebut berhasil mengidentifikasi determinan harga apartemen, sebagian besar penelitian masih berfokus pada analisis faktor secara parsial dan belum secara eksplisit mengaitkannya dengan permasalahan praktis penilaian aset, khususnya dalam konteks keterbatasan data pembandingan dan kebutuhan estimasi nilai pasar yang adaptif.

Menurut *Appraisal Institute* (2020) terdapat tiga pendekatan utama dalam penilaian properti yang dikenal luas, salah satunya adalah *sales comparison approach*. Pendekatan ini menilai properti dengan membandingkan properti sejenis yang telah terjual di area serupa, berdasarkan karakteristik seperti lokasi, ukuran, kondisi fisik, dan fitur lainnya. Penyesuaian dilakukan terhadap perbedaan karakteristik tersebut. Namun, jika properti pembandingan tidak cukup mencerminkan kesetaraan meskipun telah disesuaikan, maka properti tersebut dikeluarkan dari analisis. Akibatnya, data pembandingan yang tersisa menjadi terbatas dan berpotensi bias (*Appraisal Institute*, 2020). Untuk merespon keterbatasan ini, pendekatan berbasis data historis menjadi alternatif untuk memperkirakan harga properti secara lebih objektif (Pratama, 2023). Namun, tidak semua kawasan memiliki karakteristik yang tercermin dalam data historis, sehingga prediksi tentang menantang, terutama bagi yang kurang berpengalaman. Untuk itu, Pratama (2023) dan Tanamal et al. (2023) menyarankan penggunaan algoritma *Machine Learning* (ML) sebagai solusi yang mampu mengenali pola kompleks dan menghasilkan estimasi harga secara akurat, bahkan saat data sebelumnya terbatas. Studi lain oleh Juneja et al. (2023), Tanamal et al. (2023), Phan (2019), dan Fan et al. (2018) juga mendukung efektivitas ML dalam memprediksi harga properti.

Dalam perkembangannya, *Artificial Neural Network* (ANN) telah digunakan sejak tahun 1992 dalam penilaian properti dan terbukti unggul dalam mengolah data besar serta mendeteksi pola yang sulit diidentifikasi metode konvensional. Penelitian Prasetyo et al. (2021) di Surabaya menunjukkan ANN lebih akurat dan efisien dibanding metode konvensional. Suryawan (2022) menyebutkan bahwa model ANN juga perlu bersifat adaptif terhadap dinamika pasar. Selain itu, *Linear Regression* (LR) juga banyak digunakan untuk prediksi harga properti, dan Lathifah & Dana (2024) menyarankan evaluasi perbandingan akurasi antara LR dan ANN. Penelitian terdahulu di Surabaya umumnya masih berfokus pada prediksi harga hunian tanpa mengaitkannya dengan indikator investasi yang lazim digunakan dalam praktik penilaian, seperti *Gross Income Multiplier* (GIM) (Njo et al., 2025). Padahal, bagi penilai dan investor, konsistensi antara estimasi harga jual dan indikator berbasis pendapatan merupakan aspek penting dalam menilai kewajaran nilai pasar dan kelayakan investasi. Oleh karena itu, penelitian ini mengisi celah tersebut dengan mengintegrasikan pendekatan ML, khususnya ANN, dengan analisis GIM untuk memberikan kerangka penilaian apartemen yang lebih adaptif dan relevan terhadap praktik appraisal di Surabaya.

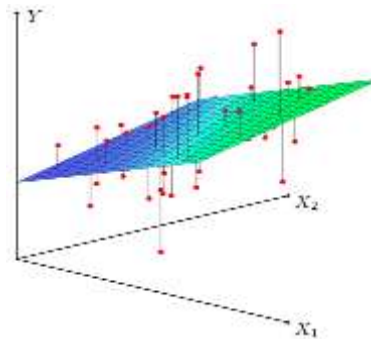
Penilaian properti bertujuan menentukan nilai pasar properti untuk keperluan transaksi, pendanaan, pelaporan keuangan, atau litigasi. Tiga konsep utama dalam penilaian yaitu nilai, biaya, dan harga. Nilai merupakan estimasi manfaat ekonomi yang bersifat subjektif. Biaya adalah total pengeluaran untuk membangun atau mengganti properti yang bersifat objektif. Harga adalah jumlah uang yang dibayarkan dalam transaksi, dapat berbeda dari nilai tergantung kondisi pasar (Ling & Archer, 2024; *Appraisal Institute*, 2020).

Penilaian properti pada dasarnya berlandaskan pada prinsip bahwa nilai ditentukan oleh manfaat ekonomi yang dirasakan dari penggunaan suatu aset, dengan mempertimbangkan berbagai aspek fisik, lingkungan, hukum, ekonomi, maupun sosial. *Appraisal Institute* (2020) menekankan

bahwa penilai mengevaluasi kondisi fisik properti, termasuk luas tanah dan bangunan, struktur, bahan bangunan, usia, pemeliharaan, topografi, drainase, utilitas (air, listrik, akses jalan), dan kualitas konstruksi. Dalam konteks apartemen, faktor struktural seperti luas unit, jumlah kamar, kualitas finishing, serta posisi lantai terbukti signifikan memengaruhi harga jual (Hoxha, et al., 2021; Museleku, 2022). Selain faktor fisik, lokasi menjadi variabel dominan dalam teori penilaian karena kedekatan dengan pusat kota, fasilitas umum, dan akses transportasi meningkatkan nilai apartemen. Studi de la Torre & Chen (2024) menunjukkan bahwa aksesibilitas dan lingkungan terbangun secara signifikan memengaruhi harga perumahan, yang sejalan dengan temuan Lee et al. (2020) bahwa pemandangan alam atau kualitas lingkungan perkotaan turut meningkatkan nilai apartemen. Faktor hukum dan politik, seperti status kepemilikan, legalitas, kepatuhan pada zonasi dan Rencana Tata Ruang Wilayah (RTRW), serta potensi pengembangan, dapat memperluas atau justru membatasi kegunaan properti, sehingga memengaruhi nilainya. Dari sisi ekonomi makro, daya beli masyarakat, tingkat suku bunga, inflasi, dan tren investasi berperan penting dalam menentukan permintaan dan harga apartemen (Ke et al., 2022). Di samping itu, faktor sosial juga tidak dapat diabaikan, misalnya perubahan demografi, preferensi gaya hidup, atau tren kerja dari rumah yang memengaruhi preferensi masyarakat terhadap tipe hunian tertentu (Khanh et al., 2024). Dengan demikian, teori penilaian apartemen menekankan interaksi antara karakteristik fisik, lokasi, hukum, ekonomi, dan sosial dalam menentukan nilai, serta menegaskan bahwa variabel-variabel tersebut harus dipahami secara kontekstual sesuai dengan dinamika pasar. Namun adanya keterbatasan data di pasar real estate Indonesia, maka penelitian ini difokuskan pada faktor fisik dan lingkungan serta hukum.

Menurut *Appraisal Institute* (2020) terdapat tiga pendekatan utama dalam penilaian properti: 1) Pendekatan Pasar (*Sales Comparison Approach*), 2) Pendekatan Pendapatan (*Income Capitalization Approach*), dan 3) Pendekatan Biaya (*Cost Approach*). Penelitian ini memfokuskan pada dua pendekatan yaitu pendekatan pasar dan pendekatan pendapatan. Pertama, pendekatan pasar berdasarkan prinsip substitusi, yaitu pembeli yang rasional tidak akan membayar lebih dari harga properti alternatif dengan utilitas serupa. Nilai pasar properti ditentukan melalui perbandingan dengan properti sejenis yang baru terjual, sedang dijual, atau sedang dalam kontrak. Pendekatan ini bersifat komparatif tanpa formula baku, melainkan menggunakan penyesuaian terhadap elemen-elemen pembanding. Nilai properti diperoleh dari “Harga Jual Pembanding $\pm$ Penyesuaian” diperoleh “Indikasi Nilai Properti Subjek”. Kedua, pendekatan pendapatan menilai properti berdasarkan kemampuannya menghasilkan pendapatan di masa depan. Terdapat dua metode yaitu *Direct Capitalization* dan *Gross Income Multiplier* (GIM). Pertama, *direct capitalization* mengubah pendapatan bersih per tahun (*Net Operating Income*) menjadi nilai properti melalui pembagian dengan tingkat kapitalisasi. Metode ini sesuai untuk properti yang stabil penghasilan bulanan atau tahunan karena menyediakan data memadai untuk menganalisis penghasilannya. Sebaliknya, untuk properti dengan arus pendapatan tidak stabil, metode kapitalisasi hasil (*Yield Capitalization*) paling sesuai sebab metode ini akan menganalisis stabilitas masa depan properti dengan mempertimbangkan pendapatan masa depan, sehingga memberikan gambaran yang lebih baik tentang potensi pertumbuhan properti (CFI team, 2025). Kedua, menggunakan faktor pengganda yaitu *Gross Income Multiplier* (GIM), rasio antara harga jual properti dengan pendapatan kotor tahunan. GIM diperoleh dari data pasar properti pembanding yang telah terjual dan disewakan, lalu diterapkan pada pendapatan kotor properti subjek untuk menghasilkan indikasi nilai. Dalam konteks apartemen di Surabaya, pendekatan pendapatan melalui *Gross Income Multiplier* (GIM) dipandang relevan sebagai indikator investasi berbasis pendapatan sewa yang relatif sederhana, khususnya ketika data pendapatan tersedia secara terbatas dan tingkat stabilitas arus kas belum sepenuhnya dapat dipastikan. Oleh karena itu, GIM digunakan dalam penelitian ini bukan sebagai alat valuasi tunggal, melainkan sebagai pembanding ekonomi untuk mengevaluasi kewajaran hasil estimasi harga jual.

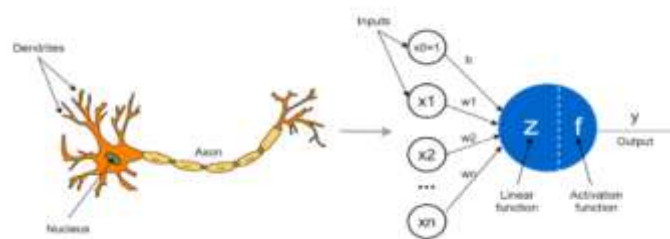
Regresi Linear adalah metode statistik yang digunakan untuk memodelkan hubungan antara variabel independen (X) dan variabel dependen (Y) melalui pendekatan garis lurus. Tujuan utamanya adalah memprediksi nilai Y berdasarkan nilai X, atau untuk memahami kekuatan hubungan di antara keduanya (Gambar 1). Dalam penelitian ini digunakan *multiple linear regression* yang memungkinkan pemodelan lebih dari satu variabel independen terhadap suatu variabel dependen.



Gambar 1. Ilustrasi pemodelan *multiple linear regression*
Sumber: Soetewey (2021, October 4).

Artificial Neural Network (ANN) banyak digunakan dalam penilaian properti karena kemampuannya menangkap hubungan non-linear dan interaksi kompleks antar variabel yang sering tidak dapat dijelaskan oleh pendekatan linier konvensional. Komponen utamanya adalah *perceptron* yang terdiri dari lima bagian: input, bobot, bias, kombinasi linier, dan fungsi aktivasi (

Gambar 2). ANN cocok untuk data dengan relasi yang tidak dapat ditangkap hanya dengan model linear.



Gambar 2. Ilustrasi *artificial neural network*
Sumber: Pramoditha (2021, December 26).

Penelitian terdahulu dilakukan oleh Kim dan Kim (2023) di Korea Selatan yang mengembangkan indeks harga jual apartemen berbasis machine learning, khususnya *Artificial Neural Network* (ANN), untuk mengatasi keterbatasan data penjualan pada suatu area dengan sedikit transaksi. Dengan ANN, indeks harga yang dihasilkan menunjukkan tingkat error rendah dengan *Mean Absolute Percentage Error* (MAPE) sebesar 4,9%. Hasil ini menegaskan bahwa ANN lebih adaptif dalam kondisi pasar dengan data terbatas dibandingkan metode regresi tradisional. Selanjutnya, ŠtubŠová et al. (2020)

menganalisis prediksi harga jual properti residensial di Kota Nitra, Republik Slovakia, dengan membandingkan regresi dan ANN. Hasil penelitian menunjukkan bahwa ANN yang dilatih menggunakan algoritma Regularisasi Bayesian (BR) dengan dua lapisan tersembunyi menghasilkan akurasi tertinggi, mengungguli model regresi dalam mengestimasi harga pasar.

Temuan serupa juga ditunjukkan oleh Zietz et al.(2008), yang menegaskan bahwa regresi hedonik cenderung kurang mampu menangkap kompleksitas hubungan non-linear antar variabel, sehingga hasilnya sering bias ketika diterapkan pada pasar dengan heterogenitas tinggi, seperti apartemen perkotaan. Sebaliknya, model ANN memiliki kemampuan untuk mempelajari pola non-linear dan interaksi variabel yang kompleks, sehingga memberikan hasil prediksi yang lebih akurat (Khanh et al, 2024). Oleh karena itu, literatur terkini menekankan pentingnya mengintegrasikan ANN dengan pendekatan konvensional sebagai cara untuk meningkatkan validitas estimasi nilai apartemen, sekaligus memperkuat dasar empiris dalam praktik penilaian properti modern.

Berdasarkan tinjauan terdahulu, hipotesis penelitian ini adalah:

- H1: Terdapat perbedaan signifikan tingkat akurasi prediksi harga jual apartemen di Surabaya antara model *Artificial Neural Network* dengan *Linear Regression*.
- H2: Tidak terdapat perbedaan signifikan antara median harga jual apartemen hasil prediksi ANN dan harga jual yang diestimasi menggunakan *gross income multiplier*.

Metode

Penelitian ini merupakan studi asosiatif kuantitatif yang bertujuan untuk menganalisis tingkat akurasi model *Machine Learning* (ML) dalam memprediksi harga jual apartemen di Surabaya, serta mengevaluasi konsistensi hasil prediksi harga jual menggunakan *Artificial Neural Network* (ANN) dengan *Gross Income Multiplier* (GIM) sebagai indikator investasi berbasis pendapatan. Data yang digunakan merupakan data sekunder *listing* unit apartemen di *platform online* dikumpulkan melalui metode *web scraping*, yang diambil dalam periode 19 September 2024 hingga 21 Oktober 2024. Periode pengumpulan data yang relatif singkat dipilih untuk menangkap kondisi pasar apartemen pada satu titik waktu (*cross-sectional snapshot*), sehingga fokus penelitian diarahkan pada evaluasi kinerja model prediksi, bukan analisis dinamika siklus pasar jangka panjang. Pendekatan ini sejalan dengan tujuan penilaian aset yang menitikberatkan pada estimasi nilai pasar pada tanggal penilaian tertentu. Secara geografis, cakupan penelitian ini mencakup *unit* apartemen di berbagai wilayah administratif Kota Surabaya, yakni Surabaya Pusat, Barat, Timur, dan Selatan. Penelitian ini menggunakan teknik *non-probability sampling* dengan pendekatan *purposive sampling*. Kriteria pemilihan sampel didasarkan pada ketersediaan atribut fisik dan lingkungan serta atribut hukum. Variabel bebas penelitian ini terdiri dari atribut fisik dan lingkungan seperti jumlah kamar mandi, kamar tidur, luas bangunan, daya listrik, arah hadap unit, status *furnished*, jarak properti serta atribut hukum dan politik seperti sertifikat kepemilikan. Variabel terikat adalah harga jual dari unit apartemen. Harga yang digunakan dalam penelitian ini merupakan harga penawaran (*listing price*) yang diperoleh dari *platform* daring. Dalam praktik penilaian properti, harga penawaran sering digunakan sebagai proksi nilai pasar ketika data transaksi aktual tidak tersedia secara terbuka. Meskipun harga *listing* berpotensi mengandung bias negosiasi, penggunaan dataset yang besar diharapkan dapat mereduksi bias individual dan menangkap kecenderungan harga pasar secara agregat.

Proses pengumpulan data dimulai dengan identifikasi *platform online* yang menyediakan *listing* apartemen seperti Brighton, Xavier Marks, Propnex Indonesia, Ray White Indonesia hingga situs lain seperti Rumah123 dan OLX. Pemilihan situs didasarkan pada kelengkapan data dan kemudian proses ekstraksi data menggunakan *web scraping*. *Web scraping* adalah teknik untuk mengekstrak data atau informasi dari situs web secara otomatis (Adila et al., 2022). Teknik ini sangat berguna ketika data yang dibutuhkan tidak tersedia secara langsung atau jumlahnya yang cukup untuk dianalisis lebih

lanjut. Selain sebagai alat pengumpulan data, *web scraping* juga memungkinkan proses ekstraksi informasi dilakukan secara otomatis dan efisien (Sembiring & Sari, 2019).

Dalam penelitian ini, data diperoleh dari situs resmi Brighton, *real estate agent company* yang memiliki banyak cabang di Indonesia, termasuk Kota Surabaya. Informasi properti tersedia cukup lengkap dan bervariasi. Proses *web scraping* diawali dengan memasukkan informasi *login* serta kata kunci pencarian apartemen. Sistem kemudian membuat *dictionary* URL berdasarkan kata kunci dan memproses setiap URL secara berurutan. Untuk setiap halaman, dilakukan permintaan ke situs dan pemeriksaan terhadap status respons. Jika permintaan berhasil, sistem akan mem-*parsing* konten halaman, mengekstrak tautan (*link*) yang memiliki atribut “*href*”, dan menyimpannya sebagai data *link* properti. Proses ini berulang hingga batas maksimum halaman tercapai (hingga halaman ke-9999) atau seluruh kata kunci telah selesai diproses. Seluruh data yang berhasil dikumpulkan diekspor dalam format excel (xlsx) untuk tahapan selanjutnya.

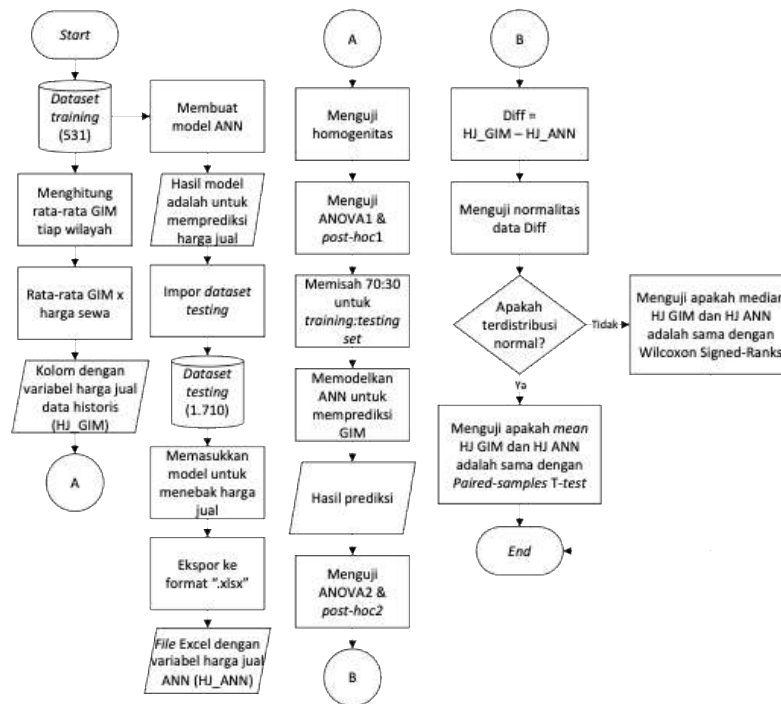
Tahap berikutnya adalah *pre-processing* data, yang bertujuan untuk membersihkan dan mempersiapkan data mentah agar siap dianalisis. Proses ini mencakup tiga tahapan utama: penghapusan duplikasi, penanganan *missing value*, dan identifikasi *outlier*. Penghapusan duplikasi dilakukan dengan menyaring *entri* yang memiliki nilai identik pada seluruh variabel menggunakan fitur “*remove duplicates*” di Microsoft Excel. Langkah ini penting agar proses pelatihan model tidak dipengaruhi oleh bias akibat keberadaan data yang berulang. Proses penghapusan data menyebabkan penyempitan rentang harga ekstrem, namun distribusi harga jual secara umum tetap mempertahankan pola asimetris yang mencerminkan karakteristik pasar apartemen Surabaya. Selanjutnya, penanganan *missing value* dilakukan dengan dua pendekatan. Pertama, variabel independen yang memiliki banyak nilai kosong dan dianggap kurang relevan dihapus dari dataset, sebagaimana pendekatan yang juga diterapkan oleh Quang et al. (2020). Kedua, metode *case-wise deletion* digunakan untuk menghapus baris data yang mengandung nilai kosong pada satu atau lebih variabel. Dengan cara ini, hanya data yang lengkap dan relevan yang digunakan dalam analisis selanjutnya.

Untuk mendeteksi *outlier*, dilakukan analisis visual menggunakan *scatter plot*. *Outlier* dikenali sebagai titik data yang berada jauh dari pola kelompok data utama. IBM SPSS Modeler 18.0 dimanfaatkan untuk mendeteksi serta memvisualisasikan *outlier*, terutama ketika nilai standar deviasi mencapai tiga kali rata-rata atau nilai ekstrim mencapai lima kali rata-rata. Metode *Inter-Quartile Range* (IQR) juga digunakan dalam identifikasi *outlier* karena bersifat *robust* terhadap distribusi data yang tidak simetris dan tetap menjaga variasi data secara keseluruhan. Pengembangan model prediksi dilakukan dengan menggunakan aplikasi IBM SPSS Modeler 18.0. Dataset utama dibagi menjadi dua bagian melalui metode *data splitting* dengan proporsi 70:30, di mana 70% digunakan sebagai *training set* dan 30% sebagai *testing set*. Pembagian ini bertujuan untuk mengevaluasi performa model, baik terhadap data yang telah dilatih maupun terhadap data baru yang belum pernah dikenali sebelumnya. Dengan demikian, dapat dipastikan bahwa model memiliki kemampuan generalisasi yang baik dan mampu melakukan prediksi secara akurat.

Pengaturan rentang parameter merupakan langkah penting dalam pengembangan model AI untuk menentukan nilai optimal yang memengaruhi kinerja model. Setiap algoritma ML memiliki parameter spesifik, seperti jumlah pohon pada *random forest* atau jumlah *neuron* pada ANN. Rentang nilai ini harus ditentukan terlebih dahulu agar pencarian kombinasi terbaik dapat dilakukan secara sistematis. Dalam penelitian ini, rentang parameter menggunakan metode *random search* berdasarkan referensi dari studi sebelumnya yang menggunakan rentang lebih luas. Metode ini dimodifikasi menjadi *local random search* dengan variasi kecil di sekitar parameter terbaik dari penelitian Kim & Kim (2023) serta Sugiarto, et al. (2025). Model ANN digunakan pada penelitian ini terdiri dari satu *hidden layer* dengan lima *neuron*, menggunakan fungsi aktivasi *non-linier* dan proses pelatihan berbasis *supervised learning*. Proses pelatihan dihentikan berdasarkan kriteria *konvergensi error* untuk menghindari *overfitting*. Model LR digunakan sebagai pembanding dengan asumsi

hubungan linier antar variabel. Evaluasi kinerja model dilakukan menggunakan empat metrik (*linear correlation, mean absolute error, mean absolute percentage error, dan root mean square error*) serta uji statistic (uji homogenitas, ANOVA, *post-hoc*, normalitas, dan *T-test*) untuk menguji hipotesis penelitian.

Gross Income Multiplier (GIM) digunakan untuk evaluasi properti saat data terbatas pada nilai jual dan sewa. Data pendapatan sewa diperoleh dari informasi pasar yang tersedia pada platform properti daring dan digunakan untuk menghitung nilai GIM historis per wilayah. Dalam penelitian ini, GIM tidak digunakan sebagai alat valuasi utama, melainkan sebagai indikator investasi berbasis pendapatan untuk mengevaluasi kewajaran hasil estimasi harga jual yang dihasilkan oleh model ANN. Pengujian dilakukan dengan uji homogenitas, *post-hoc*, dan uji normalitas, seperti pada Gambar 3.



Gambar 3. Bagan Alir Proses Analisa GIM dan Harga Jual

Uji homogenitas dan ANOVA digunakan untuk menguji perbedaan nilai GIM antar wilayah di Surabaya. Uji post-hoc dilakukan untuk mengidentifikasi pasangan wilayah yang berbeda signifikan. Uji normalitas digunakan untuk menentukan metode uji lanjutan, sedangkan *Wilcoxon Signed-Ranks Test* digunakan untuk membandingkan median harga jual hasil prediksi ANN dengan estimasi harga berbasis GIM.

Hasil dan Pembahasan

Dataset awal hasil *web scraping* terdiri dari 2.658 observasi, kemudian 1.204 data dieliminasi karena duplikasi listing unit yang sama pada beberapa agen. Selanjutnya, 612 observasi dieliminasi akibat *missing value* pada variabel kunci penilaian seperti luas bangunan, harga, atau status legalitas. Sebanyak 271 observasi lainnya dihapus pada tahap identifikasi *outlier* karena berada di luar rentang inter-kuartil yang wajar untuk pasar apartemen Surabaya. Setelah seluruh tahapan *pre-processing*, jumlah data akhir yang digunakan dalam pemodelan adalah 571 observasi. Alur penyaringan data ini dilakukan untuk memastikan hanya unit apartemen dengan atribut fisik, lingkungan, dan hukum

yang lengkap digunakan dalam analisis. Distribusi lokasi apartemen yang dijual terbanyak di “Waterplace” disusul “Puncak Dharmahusada”.

Tabel 1. Deskriptif Dataset Apartemen Jual

Variabel	Satuan	Min.	Max.	Mean	Median	Std.dev.
Luas Bangunan	m ²	18,00	98,62	44,94	40,00	19,27
Harga	Rupiah	150 juta	2.290 juta	743 juta	600 juta	482 juta
Listrik	Watt	450	6.600	1.966	2.200	803,75
Jarak ke Tol	km	1,00	12,00	5,83	4,50	2,97
Jarak ke Universitas	km	0,50	7,00	3,38	3,00	1,77
Jarak ke Mall	km	0,00	7,00	2,05	2,00	2,10
Jarak ke Juanda	km	10,00	34,00	23,86	25,00	6,16
Jarak ke Tugu Pahlawan	km	2,00	17,00	11,29	12,00	3,83

Tabel menampilkan luas unit apartemen bervariasi antara 18,00 m² hingga 98,62 m², dengan nilai rata-rata 44,94 m² dan median 40 m². Variabel harga menunjukkan rentang yang sangat lebar, mulai dari Rp.150 juta hingga Rp.2,29 miliar, dengan rata-rata Rp.744 juta dan simpangan baku Rp.483 juta, mencerminkan adanya kesenjangan harga yang cukup lebar antar unit. Konsumsi daya listrik rata-rata sebesar 1.966 W menandakan bahwa sebagian besar apartemen menggunakan daya menengah. Jarak ke berbagai fasilitas, seperti jalan tol, universitas, mal, Bandara Juanda, dan Tugu Pahlawan memiliki simpangan baku yang relatif rendah dibandingkan rentang maksimum, mengindikasikan bahwa persebaran jarak antar sampel tidak terlalu ekstrem.

Tabel 2. Frekuensi Dataset Apartemen Jual

Variabel	N	N Kumulatif	%	% Kumulatif
Kamar	1	234	234	41,00%
Tidur	2	295	529	51,70%
	3	36	565	6,30%
	4	6	571	1,10%
Kamar	1	495	495	86,70%
Mandi	2	70	565	12,30%
	3	6	571	1,10%
Sertifikat	SHM	91	91	15,90%
	HGB	65	156	11,40%
	PPJB	357	513	62,50%
	Lain-lain	58	571	10,20%
Hadap	Barat	43	43	7,50%
	Barat Daya	10	53	1,80%
	Barat Laut	10	63	1,80%
	Selatan	153	216	26,80%
	Tenggara	14	230	2,5%
	Timur	107	337	18,70%
	Timur Laut	21	358	3,70%
	Utara	213	571	37,30%
Group	Primary	2	2	0,40%
	Secondary	569	571	99,60%
Furnish	Fully-furnished	480	480	84,10%
	Non-furnished	57	537	10,00%
	Semi-furnished	34	571	6,00%

Tabel 2 menyajikan distribusi frekuensi variabel pada *dataset* apartemen yang dijual. Mayoritas apartemen memiliki dua kamar tidur (51,70%) dan satu kamar mandi (86,70%), yang mencerminkan dominasi tipe unit yang umum di pasar. Dari aspek legalitas, mayoritas kepemilikan adalah PPJB (62,5%), sedangkan sisanya sertifikat kepemilikan SHM dan HGB. Orientasi bangunan paling umum adalah menghadap Utara (37,3%) dan Selatan (26,8%). Selain itu, mayoritas unit termasuk kategori pasar sekunder (99,6%) serta dalam kondisi *fully-furnished* (84,1%). Informasi ini memberikan gambaran tentang tren pasar dan preferensi umum dalam penawaran apartemen yang tersedia.

Model *Linear Regression* digunakan sebagai model pembanding (*baseline*) untuk evaluasi kinerja relatif *Artificial Neural Network*. Oleh karena itu, analisis regresi dalam penelitian ini difokuskan pada kemampuan prediksi secara agregat, bukan pada inferensi kausal individual setiap koefisien.

Pengolahan data menggunakan teknik LR menghasilkan persamaan:

$$y = -710.116.061,1 - 203.831.218,7 \times \text{kamar tidur} + 157.093.251,9 \times \text{kamar mandi} + 21.427.815,2 \times \text{luas bangunan} + 17.953,2 \times \text{daya listrik} + 48.399.538,1 \times \text{jarak ke tol} + 84.817.774,8 \times \text{jarak ke universitas} - 11.177.491,3 \times \text{jarak ke mall} + 8.667.087,8 \times \text{jarak ke juanda} - 11.133.548,9 \times \text{jarak ke Tugu Pahlawan} \dots\dots\dots (1)$$

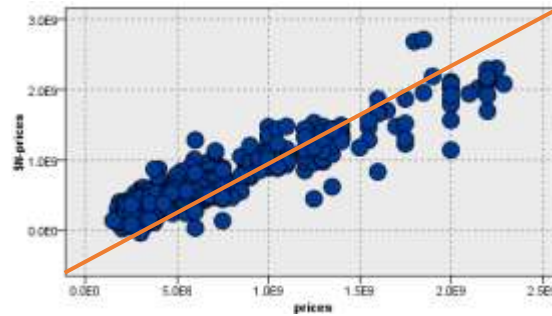
Hasil dari persamaan regresi linear menggambarkan hubungan antara harga properti dan sejumlah variabel penjelas, baik secara internal maupun eksternal. Nilai konstanta menunjukkan harga dasar ketika seluruh variabel bernilai nol, meskipun secara praktis lebih berfungsi sebagai penyesuaian model. Koefisien regresi masing-masing variabel memberikan arah dan besarnya pengaruh, misalnya jumlah kamar mandi dan luas bangunan yang berpengaruh positif, sehingga semakin banyak fasilitas dan semakin luas bangunan maka harga cenderung meningkat. Sebaliknya, jumlah kamar tidur justru berpengaruh negatif, yang mencerminkan karakteristik khusus dari data properti apartemen yang tidak bisa bertambah disbanding rumah tinggal. Faktor eksternal seperti kedekatan dengan mall dan landmark kota berkontribusi menaikkan harga, sedangkan jarak yang lebih jauh dari tol, universitas, atau bandara dalam model ini justru diasosiasikan dengan harga yang lebih tinggi, yang dikaitkan dengan preferensi masyarakat terhadap lingkungan yang lebih tenang dan nyaman. Persamaan ini menegaskan kombinasi faktor fisik dan aksesibilitas berperan dalam menentukan variasi harga properti.

Evaluasi kelayakan model LR dilakukan berdasarkan kinerja prediktif (MAE, RMSE, MAPE) dan konsistensi antara data training dan testing. Tabel 3 merangkum kinerja model LR pada skenario pasar yaitu apartemen jual. Pada *dataset* jual, nilai korelasi R sedikit lebih tinggi sebesar 0,796 (data *training*) dan 0,781 (data *testing*). Namun, skala harga jauh lebih besar menyebabkan MAE meningkat menjadi Rp.221–222 juta dan RMSE mendekati Rp.290 juta, MAPE sebesar 36%, menunjukkan variabilitas harga jual lebih sulit ditangkap oleh model linier sederhana. Meski demikian, konsistensi nilai metrik antara data *training* dan *testing* di kedua skenario mengindikasikan generalisasi model yang memadai.

Tabel 3. Hasil Evaluasi Model LR

Metode Evaluasi	Dataset Jual	
	Training	Testing
R	0,796	0,781
MAE (Rp)	221.006.385	222.370.558
MAPE	35,015%	37,164%
RMSE (Rp)	298.629.059	286.079.922

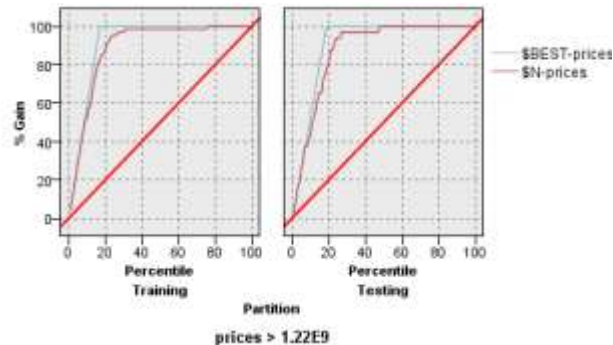
Gambar 4 membandingkan harga apartemen aktual (sumbu-X) dengan harga yang diprediksi model ANN (sumbu-Y) pada data uji. Garis oranye 45° merepresentasikan prediksi sempurna, dimana titik yang menempel pada garis berarti model menebak tepat.



Gambar 4. Komparasi Harga Jual Aktual terhadap Harga Prediksi Apartemen Model ANN

Gambar 5 menunjukkan model ANN mampu secara efektif mengidentifikasi apartemen dengan harga di atas Rp.5,13 juta dan Rp.1,22 miliar. Hasil analisis menunjukkan bahwa sekitar 90% kategori harga apartemen tersebut dapat terdeteksi hanya dengan memeriksa 20% data teratas yang memiliki skor prediksi tertinggi, baik pada data *training* maupun *testing*. Kesamaan pola antara kedua jenis data ini mengindikasikan kemampuan generalisasi yang kuat dengan risiko *overfitting* yang minimal. *Gains Chart* pada Gambar 5 menunjukkan evaluasi kinerja model *Artificial Neural Network* (ANN) ada dataset harga jual properti. Garis merah diagonal merepresentasikan baseline atau kondisi acak, sedangkan kurva hasil model menggambarkan seberapa jauh ANN mampu mengungguli prediksi acak tersebut. Terlihat bahwa kurva pada data *training* maupun *testing* naik dengan sangat cepat dan mencapai lebih dari 80 hingga mendekati 100 persen pada persentil awal, yang berarti sebagian besar properti dengan harga di atas 1,22 miliar dapat diidentifikasi hanya

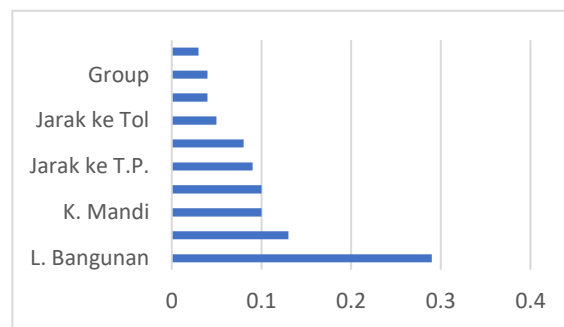
dengan sebagian kecil data. Pola yang konsisten antara grafik training dan testing juga menunjukkan bahwa model tidak mengalami *overfitting* yang berarti. Dengan demikian,



Gambar 5. Gains Ch

Gains Chart dalam penelitian ini tidak digunakan sebagai metrik utama evaluasi model regresi, melainkan sebagai alat eksploratif untuk menggambarkan kemampuan model ANN dalam mengidentifikasi unit apartemen dengan nilai harga relatif tinggi berdasarkan skor prediksi kontinu. Evaluasi utama kinerja ANN tetap didasarkan pada metrik regresi seperti MAE, RMSE, dan MAPE.

Error! Reference source not found. menunjukkan bobot kepentingan setiap variabel pada hasil pemodelan ANN untuk *dataset* penjualan apartemen. Luas bangunan menempati posisi paling dominan, dengan kontribusi 29% terhadap akurasi prediksi, menegaskan bahwa ukuran fisik unit merupakan faktor penentu utama dalam pembentukan nilai pasar. Selain itu, variabel lain yang juga berperan penting adalah jarak ke universitas, jumlah kamar mandi, jarak ke Tugu Pahlawan, dan jarak ke mal. Temuan ini mengindikasikan bahwa pembeli tidak hanya dipengaruhi oleh utilitas internal apartemen, tetapi juga oleh faktor eksternal yang berkaitan dengan aksesibilitas dan nilai simbolik atau historis dari lokasi.



Gambar 6. Derajat Kepentingan ANN sesuai Faktor Fisik dan Lingkungan

Tabel 4. menyajikan hasil pemodelan ANN untuk *dataset* penjualan menggunakan satu *hidden-layer* dengan lima *neuron*. Nilai korelasi linear (R) menunjukkan hubungan positif yang kuat, yaitu 0,931 (*training*) dan 0,899 (*testing*). Selisih yang tipis antara keduanya mengindikasikan kemampuan generalisasi yang baik tanpa indikasi *overfitting* yang serius. Secara absolut, akurasi moneterinya tergambar dari nilai MAE dan RMSE, di mana rata-rata kesalahan untuk data penjualan lebih besar (Rp 123–202 juta) karena skala harga memang lebih tinggi. Nilai MAPE memperkuat temuan ini, dengan kesalahan relatif 20–25% yang masih berada dalam batas wajar untuk aplikasi penilaian properti yang kompleks.

Tabel 4. Hasil Evaluasi Model ANN

Metode Evaluasi	Dataset Jual
-----------------	--------------

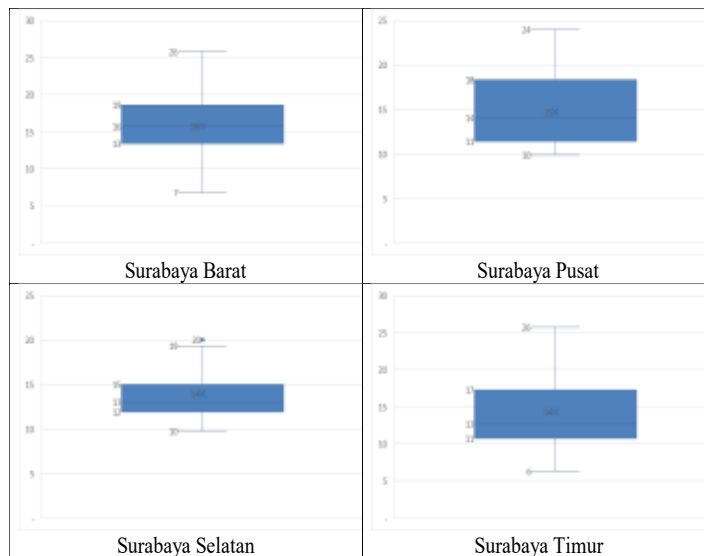
	<i>Training</i>	<i>Testing</i>
R	0,931	0,899
MAE (Rp)	123.050.314	151.807.664
MAPE	19,949%	24,624%
RMSE (Rp)	180.767.844	201.916.958

Tabel 5 menunjukkan analisis *Gross Income Multiplier* (GIM) pada empat wilayah di Kota Surabaya. Wilayah Surabaya Barat memiliki data pembanding jumlah apartemen terbanyak, diikuti Surabaya Timur, Surabaya Pusat, dan Surabaya Selatan. Nilai GIM di Surabaya Barat juga lebih tinggi dibanding wilayah lain di Surabaya,

Tabel 5. Perhitungan GIM Tiap Wilayah di Kota Surabaya

GIM/ Wilayah	Barat	Timur	Pusat	Selatan
Jumlah <i>Unit</i>	296	164	34	38
GIM	15,762	14,293	14,715	13,982

Gambar 7 menunjukkan sebaran data apartemen dengan nilai GIM di setiap wilayah Kota Surabaya. Terlihat Surabaya Pusat memiliki GIM dengan sebaran lebih besar dibandingkan Surabaya Barat dan Surabaya Timur, sedangkan Surabaya Selatan, memiliki sebaran GIM paling kecil. Surabaya Barat dan Surabaya Pusat cenderung memiliki GIM lebih tinggi dikarenakan fasilitas dan aksesibilitas pada lingkungan sekitar apartemen, sedangkan Surabaya Selatan dan Timur relatif lebih rendah disebabkan pasar apartemen masih berkembang.



Gambar 7. Persebaran GIM di Setiap Wilayah Apartemen Surabaya

Error! Reference source not found. menunjukkan hasil uji homogenitas yang dilanjutkan dengan uji ANOVA. Hasil pengujian menunjukkan bahwa *p-value* yang didapat lebih besar dari tingkat signifikansi konvensional (5%) sehingga secara statistik tidak diperlukan uji *post-hoc*. Namun, karena tujuan penelitian ini bersifat eksploratif dan menggunakan jumlah sampel yang terbatas, maka pelanggaran tingkat signifikansi hingga 10% dianggap wajar untuk mendeteksi temuan yang berpotensi signifikan (Lee, 2025). Dengan menaikkan tingkat signifikansi ANOVA menjadi 10%, maka uji *post-hoc* diperlukan. Hasil uji pertama menunjukkan adanya perbedaan nilai GIM yang signifikan antara wilayah Surabaya Barat dan Surabaya Timur. Sesuai hasil pengujian *post-hoc* pertama dengan metode *Tukey HSD* pada Gambar 8, diperoleh nilai signifikansi sebesar 0,054 antara wilayah Barat dengan Timur, yang berada di bawah batas signifikansi 10% ($\alpha = 0,10$). Hasil ini menunjukkan adanya perbedaan yang signifikan antara rata-rata nilai GIM kedua wilayah tersebut. Temuan ini mengindikasikan adanya perbedaan karakteristik ekonomi, kondisi pasar properti, atau faktor lainnya antara wilayah Barat dan Timur secara statistik.

Tabel 6. Test of Homogeneity of Variances

		Levene Statistics	df1	df2	Sig
Jual/Sewa	Based on Mean	2.544	3	528	.055
	Based on Median	1.758	3	528	.154
	Based on Median and with adjusted of	1.758	3	346,824	.155
	Based on trimmed mean	2.135	3	528	.095

Tabel 7. ANOVA

Jual/Sewa	Sum of Squares	df	Mean Square	F	Sig
Between Groups	251.984	3	83.995	2.587	.052
Within Groups	17141.857	528	32.466		
Total	17393.840	531			

Tabel 8. Post Hoc Tests Multiple Comparisons

Dependent Variabel: Jual/Sewa
Tukey HSD

(I) Wilayah Num	(J) Wilayah Num	Mean Difference (I-J)	Std Error	Sig	90% Confidence Interval	
					Lower Bound	Upper Bound
Barat	Pusat	.7815130575	1.031771729	.873	-1.58847314	3.151499259
	Selatan	1.567136875	.98185554329	.382	-.688191277	3.822465028
	Timur	1.41436180*	.5546552196	.054	.1403152482	2.688408358
Pusat	Barat	-.781513057	1.031771729	.873	-3.15149926	1.588473144
	Selatan	.7856238178	1.345076420	.937	-2.30402524	3.875272880
	Timur	.6328487457	1.073700562	.935	-1.83344825	3.099145743
Selatan	Barat	-1.56713688	.9818554329	.382	-3.82246503	.6881912774
	Pusat	-.785623818	1.345076420	.937	-3.87527288	2.304025245
	Timur	-.152775072	1.025826539	.999	-2.50910513	2.203554990
Timur	Barat	-1.4143618*	.5546552196	.054	-2.68840836	-.140315248
	Pusat	-.632848746	1.073700562	.935	-3.09914574	1.833448251
	Selatan	.1527750721	1.025826539	.999	-2.20355499	2.509105134

Tabel 9. Hasil Uji Homogenitas dan ANOVA

Jual/Sewa	Sum of Squares	df	Mean Square	F	Sig
Between Groups	539.615	3	179.872	7.593	.000
Within Groups	12507.765	528	23.689		
Total	13047.380	531			

Tabel 10. Post Hoc Tests Multiple Comparisons

Dependent Variable: Jual/Sewa
Tukey HSD

(I) Wilayah Num	(J) Wilayah Num	Mean Difference (I-J)	Std Error	Sig	90% Confidence Interval	
					Lower Bound	Upper Bound
Barat	Pusat	-.401433267	.8813421692	.969	-2.67285566	1.869989121
	Selatan	1.566311824	.8387035356	.243	-.595229954	3.727844603
	Timur	2.10883647*	.4737879713	.000	.8877754121	3.329897537
Pusat	Barat	.4014332673	.8813421692	.969	-1.86998912	2.672855655
	Selatan	1.967745092	1.148967874	.318	-.993410568	4.928900751
	Timur	2.51026974*	.9171578907	.032	.1465419647	4.873997519
Selatan	Barat	-1.56631182	.8387035356	.243	-3.72784460	.5952209542
	Pusat	-1.96774509	1.148967874	.318	-4.92890075	.9934105677
	Timur	.5425246501	.8762637719	.926	-1.71580953	2.800858835
Timur	Barat	-2.10883647*	.4737879713	.000	-3.32989754	-.887775412
	Pusat	-2.51026974*	.9171578907	.032	-4.87399752	-.146541965
	Selatan	-.542524650	.8762637719	.926	-2.80085883	1.715809535

*The mean difference is significant at the 0.05 level

Selanjutnya, dilakukan uji ANOVA kedua yang menghasilkan nilai signifikan, sehingga dianalisis lebih lanjut melalui uji *post-hoc2*. Hasil uji *post-hoc2* menggunakan metode *Tukey HSD* (Gambar 9) menunjukkan bahwa wilayah Barat secara signifikan berbeda dengan wilayah Timur ($p = 0,000$), dan wilayah Pusat juga berbeda signifikan dengan wilayah Timur ($p = 0,032$). Temuan ini

mengindikasikan bahwa hasil prediksi harga jual apartemen dengan model ANN di wilayah Barat dan Pusat secara statistik berbeda dengan wilayah Timur. Pola ini konsisten dengan hasil uji *post-hoc* pertama yang mengindikasikan bahwa pendekatan penilaian menggunakan ANN mampu mengungkap pola geografis dalam perbedaan nilai GIM.

Pada tahapan uji normalitas (Gambar 10) menunjukkan data tidak terdistribusi normal, sehingga analisis dilanjutkan menggunakan *Wilcoxon Signed-Ranks Test*. Hasil uji menunjukkan nilai signifikansi sebesar 0,110 lebih besar dari α (0,05). Hal ini mengindikasikan tidak terdapat perbedaan yang signifikan antara median harga jual prediksi ANN dan harga jual berdasarkan GIM historis. Dengan demikian, model ANN dinilai mampu menghasilkan estimasi harga jual yang sebanding dengan pendekatan GIM yang dihitung melalui perkalian harga sewa dan GIM rata-rata wilayah. Temuan ini menegaskan bahwa kedua pendekatan relatif sepadan dan dapat digunakan secara komplementer dalam konteks penelitian ini.

Tabel 11. Test of Normality

Table 11: Test of Normality						
	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig	Statistic	df	Sig
HJ-Diff	.111	1710	.000	.930	1710	.000

^a Lilliefors Significance Correction

Tabel 12. Wilcoxon Signed Ranks Test Ranks

		N	Mean Rank	Sum of Ranks
HJ-ANN -HJ-GIM	Negative Ranks	916 ^a	762.93	698847.00
	Positive Ranks	794 ^b	962.29	764058.00
	Ties	0 ^c		
	Total	1710		

^a $HJ_ANN < HJ_GIM$

^b $HJ_ANN > HJ_GIM$

^c $HJ_ANN = HJ_GIM$

Tabel 13. Hasil Uji Wilcoxon Signed-Ranks

	HJ_ANN – HJ_GIM
Z	-1.597 ^b
Asymp Sig (2-tailed)	.110

^a Wilcoxon Signed Ranks Test

^b Based on negative ranks

Jumlah observasi pada uji Wilcoxon (N = 1.710) merupakan hasil *pairing* antara harga jual hasil prediksi ANN dengan estimasi harga berbasis GIM pada level unit-wilayah, maka satu unit apartemen dapat berkontribusi lebih dari satu pasangan observasi dalam analisis. Dengan demikian, jumlah observasi pada uji Wilcoxon tidak merepresentasikan jumlah unit unik, melainkan jumlah pasangan perbandingan yang dihasilkan dari struktur data wilayah.

Pembahasan

Salah satu temuan utama penelitian ini adalah keunggulan model ANN dibandingkan LR dalam memprediksi harga jual apartemen di Surabaya. ANN menunjukkan nilai korelasi (R) tertinggi sebesar 0,931 pada *dataset* jual (*training*) serta menghasilkan MAPE terendah, menegaskan kemampuannya menangkap pola hubungan yang kompleks dan *non-linear* antara atribut apartemen dengan harga jual. Hasil ini sejalan dengan studi Antipov dan Pokryshevskaya (2012) serta Kok et al. (2017) yang juga menegaskan superioritas pendekatan machine learning dalam mengolah data

properti dengan struktur variabel yang saling berinteraksi. Sebaliknya, LR tetap menunjukkan adanya hubungan linier yang cukup kuat namun keterbatasannya terlihat dalam menjelaskan interaksi variabel yang tidak sepenuhnya linear, sedangkan fleksibilitas arsitektur ANN memungkinkan penyesuaian lebih efektif terhadap karakteristik data.

Konsistensi variabel luas bangunan sebagai prediktor utama harga jual, baik pada ANN maupun LR, memperkuat prinsip fundamental valuasi properti yang berlaku universal. Temuan ini selaras dengan studi Özdemir et al. (2022) serta Lee et al. (2020) yang menunjukkan bahwa atribut fisik seperti luas bangunan, jumlah kamar tidur dan jumlah kamar mandi merupakan determinan penting dalam pembentukan nilai pasar hunian. Lebih lanjut, hasil LR dalam penelitian ini mengungkapk hubungan menarik terkait variabel jarak ke universitas, di mana semakin jauh dari universitas, harga jual cenderung lebih tinggi. Hal ini dapat diinterpretasikan bahwa apartemen premium di Surabaya lebih banyak berlokasi di kawasan yang relatif jauh dari lingkungan kampus, seperti kasus unit termahal di Grand Sungkono Lagoon Tower Caspian yang mencapai Rp.2,29 miliar dengan jarak tujuh kilometer dari universitas terdekat. Fenomena serupa ditampilkan pada studi Wang (2023) yang menemukan rumah berjarak dalam radius 0,5 miles (setara 0,8 km) dari universitas mengalami penurunan harga paling maksimal sekitar 7% karena faktor kepadatan, kebisingan, dan karakteristik penyewa mahasiswa. Di sisi lain, jumlah kamar mandi lebih banyak dipersepsikan sebagai simbol kenyamanan dan nilai tambah jangka panjang, sehingga mendorong harga jual lebih tinggi, sejalan dengan teori *hedonic pricing*.

Temuan lain yang signifikan adalah kemampuan ANN dalam menghubungkan prediksi harga jual dengan analisis *Gross Income Multiplier* (GIM). ANN tidak hanya memberikan estimasi harga jual yang akurat, tetapi juga mampu menghasilkan prediksi GIM yang konsisten dengan data historis dan signifikan secara statistik. Dengan demikian, ANN terbukti unggul tidak hanya dalam akurasi prediksi harga, tetapi juga dalam menilai daya tarik investasi antarwilayah di Surabaya. Sebaliknya, LR cenderung terbatas dalam menangkap variasi GIM karena ketergantungannya pada asumsi linearitas. Dengan demikian, penelitian ini tidak hanya menegaskan superioritas ANN dalam akurasi prediksi harga sebagaimana ditunjukkan oleh Kim & Kim (2023), Buranasiri & Laokulrach (2025), serta ŠtubŠová et al. (2020), tetapi juga memperluas kerangka analisis dengan mengintegrasikan indikator investasi seperti GIM, yang jarang diangkat dalam studi prediksi properti. Hal ini menjadi nilai tambah penelitian sekaligus memberikan kontribusi praktis bagi investor dan pengembangan dalam menilai daya tarik pasar apartemen di Surabaya.

Kesimpulan

Hasil analisis menunjukkan bahwa model *Artificial Neural Network* (ANN) memiliki kinerja prediksi yang relative lebih baik dibandingkan *Linear Regression* dalam konteks data yang digunakan. Namun, nilai MAE dan RMSE yang masih besar secara absolut serta adanya deviasi pada segmen harga tinggi mengindikasikan bahwa akurasi ANN masih memiliki keterbatasan, khususnya untuk estimasi nilai unit apartemen dengan harga ekstrem.

Analisis perbandingan antara harga jual hasil prediksi ANN dan estimasi harga berbasis GIM menunjukkan tidak terdapat perbedaan median yang signifikan secara statistik. Temuan ini mengindikasikan adanya konsistensi deskriptif antara kedua pendekatan dalam konteks data penelitian ini, tanpa memberikan bukti inferensial yang kuat mengenai perbedaan atau keunggulan salah satu pendekatan. Hasil uji ANOVA mengonfirmasi bahwa prediksi GIM oleh ANN selaras dengan pola historis, di mana wilayah dengan nilai GIM lebih rendah menunjukkan perbedaan karakteristik pasar secara deskriptif, yang dapat menjadi indikasi awal bagi analisis investasi lanjutan, namun tidak dapat disimpulkan sebagai ukuran langsung kecepatan pengembalian modal dalam penelitian ini. Secara keseluruhan, penelitian ini menegaskan potensi ANN sebagai alat valuasi properti yang tidak hanya akurat dalam memproyeksikan harga jual apartemen. Dalam penelitian ini, integrasi ANN dengan indikator GIM bersifat eksploratif dan digunakan sebagai alat pembanding untuk

mengevaluasi kewajaran hasil estimasi harga jual. Oleh karena itu, temuan penelitian ini belum dapat dijadikan dasar tunggal untuk pengambilan keputusan investasi, melainkan sebagai masukan awal yang memerlukan pengujian lanjutan dengan data dan metode yang lebih komprehensif. Penelitian selanjutnya disarankan untuk memasukkan variabel tambahan yang mencakup aspek non-fisik, seperti kualitas pengelolaan apartemen, tingkat okupansi, serta kebijakan tata kota, guna memberikan gambaran yang lebih komprehensif mengenai faktor-faktor yang memengaruhi harga jual properti. Selain itu, integrasi metode penilaian operasional—misalnya biaya pemeliharaan, tingkat kekosongan unit, dan indikator ekonomi makro—dapat memperkaya analisis, khususnya dalam mengevaluasi *Gross Income Multiplier* (GIM) secara lebih menyeluruh. Penggunaan dataset yang lebih luas dan bersumber dari berbagai platform properti juga direkomendasikan untuk meningkatkan generalisasi serta validitas model prediksi. Upaya ini sekaligus dapat mengatasi ketidakseimbangan jumlah data antarwilayah, sehingga penilaian GIM menghasilkan temuan yang lebih konsisten dan meyakinkan. Penelitian ini memiliki sejumlah keterbatasan yang perlu dicermati dalam menafsirkan hasilnya. Penyusutan data yang signifikan akibat proses pembersihan, ketidakseimbangan jumlah unit antar wilayah, penggunaan harga penawaran sebagai proksi nilai pasar, serta keterbatasan data pendapatan sewa memengaruhi kekuatan inferensi statistik dan generalisasi temuan.

Penelitian ini memberikan kontribusi empiris terbatas pada literatur penilaian properti dengan menunjukkan potensi penggunaan Artificial Neural Network sebagai alat bantu prediksi harga jual apartemen dalam kondisi data pasar yang terbatas. Integrasi indikator Gross Income Multiplier dalam penelitian ini masih bersifat eksploratif dan belum membentuk kerangka konseptual penilaian yang inferensial, sehingga temuan ini lebih tepat diposisikan sebagai dasar pengembangan penelitian lanjutan. Model ANN dapat dijadikan alat bantu dalam menilai harga jual apartemen secara lebih akurat, terutama pada kondisi pasar yang kompleks dengan interaksi variabel non-linear. Analisis berbasis GIM yang diintegrasikan ke dalam ANN memberikan perspektif tambahan bagi penilai dalam mengevaluasi potensi investasi, sehingga tidak hanya berhenti pada estimasi harga, tetapi juga menilai kelayakan finansial properti. Hasil penelitian ini dapat menjadi acuan dalam menyusun kebijakan tata kota dan perencanaan infrastruktur, terutama terkait lokasi apartemen premium yang cenderung berkembang di area jauh dari kampus atau pusat keramaian mahasiswa. Analisis GIM antarwilayah dapat membantu pemerintah memahami disparitas potensi investasi di Surabaya, sehingga kebijakan insentif atau regulasi dapat diarahkan secara lebih tepat untuk menyeimbangkan perkembangan pasar properti antarwilayah.

Daftar Pustaka

- Adila, N., Sembiring, F., & Jatmiko, W. (2022). Implementation of web scraping for journal data collection on the SINTA website. *Sinkron: Jurnal dan Penelitian Teknik Informatika*, 6(4), 2478-2485. doi:10.33395/sinkron.v7i4.11576
- Anastasia, N., Sugiarto, H. D., & Prayogo, D. (2025). Artificial intelligence algorithms to predict housing market prices in Surabaya, Indonesia. *International Journal of Housing Markets and Analysis*, 19(4), 1053-1077. <https://doi.org/10.1108/IJHMA-01-2025-0022>
- Appraisal Institute. (2020). *The appraisal of real estate* (15th ed.). Chicago, IL: Appraisal Institute.
- Buranasiri, J., & Laokulrach, M. (2025). Appropriate rental price prediction for condominiums in Pattaya, Thailand using artificial neural network approach. *Journal of Social Economics Research*, 12(3), 161-175. <https://doi.org/10.20944/preprints202504.0737.v1>
- Bedi, A., & Rahadi, R. A. (2024). Adapting to change: A study of post-pandemic millennial housing preferences. *Ekombis Review: Jurnal Ilmiah Ekonomi dan Bisnis*, 12(2), 1733-1744. doi:<https://doi.org/10.37676/ekombis.v12i2>

- CFI Team (2025). Direct capitalization. method. CFI Education Inc. Retrieved from: <https://corporatefinanceinstitute.com/resources/commercial-real-estate/direct-capitalization-method/>
- Colliers Indonesia. (2024). *Residential services forecast report*. Jakarta: Colliers.
- Darmawan, A. D. (2024, 12 18). *Data BPS 2024, 21,47% Penduduk kota Surabaya masih anak-anak*. Databoks 18 Desember 2024. <https://databoks.katadata.co.id/demografi/statistik/c97354cd852ed9e/data-bps-2024-21-47-penduduk-kota-surabaya-masih-anak-anak>
- de la Torre, V. G. F. M., & Chen, X. (2024). Housing price determinants in Ecuador: A spatial hedonic analysis. *International Journal of Housing Markets and Analysis*, 17(6), 1461-1487. <https://doi.org/10.1108/IJHMA-09-2023-0121>
- Fan, C., Cui, Z., & Zhong, X. (2018). House prices prediction with machine learning algorithms. *10th International Conference on Machine Learning and Computing*, 6-10. doi:<https://doi.org/10.1145/3195106.3195133>
- Hoxha, V., Hoxha, D., & Hoxha, J. (2021). Study of factors influencing apartment prices in Prishtina, Kosovo. *International Journal of Housing Markets and Analysis*, 14(6), 1151-1168. <https://doi.org/10.1108/IJHMA-10-2021-0086>
- Juneja, S., Chaudhary, N., Gupta, R., Kaushik, O., Ishan, M., & Sharma, A. (2023). House price prediction using machine learning algorithms. *International Journal for Research in Applied Science and Engineering Technology*, 11(6), 3156-3164. <https://doi.org/10.22214/ijraset.2023.54259>
- Ke, X., Yang, C., Shi, W., Mougharbel, A., Guo, H., & Zheng, M.. (2022). Impact of different ecological landscapes on housing prices empirical evidence from wuhan through the hedonic pricing model appraisal. *Journal of Housing and the Built Environment*, 38, 1289-1308. <https://doi.org/10.1007/s10901-022-09990-w>
- Khanh, D. V., Son, T. T., Tuan, P. A. & Thang, N. V. (2025). Empirical analysis using a hedonic model of high-end apartment pricing. *International Journal of Engineering*, 38(10), 2426-2434. <https://doi.org/10.5829/ije.2025.38.10a.18>
- Kim, J., & Kim, K. M. (2023). A study on the machine learning-based apartment price index. *Journal of Korea Planning Association*, 58(4), 160-177. <https://doi.org/10.17208/jkpa.2023.08.58.4.160>
- Lathifah, U., & Dana, R. D. (2024). Implementasi metode linear regression untuk prediksi harga properti real estate menggunakan rapidminer. *JATI: Jurnal Mahasiswa Teknik Informatika*, 8(1), 1129-1137. <https://doi.org/10.36040/jati.v8i1.8919>
- Lauren, C. C. (2023). Implikasi hukum penipuan properti dalam industri konstruksi di Indonesia. *Jurnal Hukum dan HAM Wara Sains*, 2(5). 389-400. <https://doi.org/10.58812/jhhws.v2i05.330>
- Lee, H., Lee, B., & Lee, S. (2020). The Unequal Impact of Natural Landscape Views on Housing Prices: Applying Visual Perception Model and Quantile Regression to Apartments in Seoul. *Sustainability*, 12(19), 1-19. <https://doi.org/10.3390/su12198275>
- Lee, S. (2025, May 18). *Mastering alpha in AP statistics*. Dipetik June 6, 2025, dari Number Analytics: <https://www.numberanalytics.com/blog/alpha-selection-strategies-ap-stats#criteria-for-choosing-alpha>
- Ling, D. C., & Archer, W. R. (2024). *Real estate principles: A value approach*. New York: McGraw Hill.
- Museleku, E. K. (2022). Modelling apartments values in the Nairobi metropolitan area, Kenya. *Property Management*, 41(2), 212-227. <https://doi.org/10.1108/PM-03-2022-0023>
- Ong, S. E., Petrova, M., & Spieler, A. (2013). Demand for university student housing: An empirical analysis. *Journal of Housing Research*, 22(2), 141-164. <https://www.jstor.org/stable/24862547>

- Özdemir, M., Yıldız, K., & Büyüktanır, B. (2022). Housing price estimation with deep learning: A case study of Sakarya Turkey. *BSEU Journal of Science*, 9(1), 138-151. <https://doi.org/10.35193/bseufbd.998331>
- Phan, T. D. (2019). Housing price prediction using machine learning algorithms: The case of Melbourne city, Australia. *Proceedings of the 2018 International Conference on Machine Learning and Data Engineering*, 8, 35-42. <https://doi.org/10.1109/iCMLDE.2018.00017>
- Pramoditha, R. (2021, December 26). The concept of artificial neurons (perceptrons) in neural networks. Retrieved April 2025, from Towards data science: <https://towardsdatascience.com/the-concept-of-artificial-neurons-perceptrons-in-neural-networks-fab22249cbfc/>
- Prasetyo, B., Nugroho, H., & Wicaksono, T. (2021). Application of artificial neural network in property valuation: Case study in Surabaya. *Jurnal Teknologi Informasi*, 15(1), 45-57.
- Pratama, I. P. (2023). Implementasi support vector machine dalam prediksi harga rumah. *JAVOK: Jurnal Akademisi Vokasi*, 2(2), 101-113. <https://doi.org/10.63604/javok.v2i2.54>
- Quang Truong, Minh D. Nguyen, H. Dang, B. Mei, dan rekan. 2020. Housing Price Prediction via Improved Machine Learning Techniques. *Procedia Computer Science*, 174, 433-442. <https://doi.org/10.1016/j.procs.2020.06.111>
- Sembiring, F., & Sari, D. P. (2019). Design process data storage and organize data scraping. *INTEGRATED: Information Technology and Vocational Education*, 1(1). 19-22. <https://doi.org/10.17509/integrated.v1i1.16837>
- Snapcart. (2025). *The decline of the middle class in Indonesia in 2025 and its impact on the economy and business*. diambil dari Snapcart: <https://snapcart.global/the-decline-of-the-middle-class-in-indonesia-in-2025-and-its-impact-on-the-economy-and-business/>
- Soetewey, A. (2021). *Multiple linear regression made simple*. Dipetik April 2025, dari Stats and R: <https://statsandr.com/blog/multiple-linear-regression-made-simple/>
- ŠtubŠová, M., Urbaníková, M., Hudáková, J., & Papcunová, V. (2020). Estimation of Residential Property Market Price: Comparison of Artificial Neural Networks and Hedonic Pricing Model. *Emerging Science Journal*, 4(6), 530-538. <https://doi.org/10.28991/esj-2020-01250>
- Sugiarto, H. D., Prayogo, D., & Njo, A. (2025, January 22). Pengembangan algoritma prediktif untuk harga rumah tinggal di Surabaya berbasis machine learning menggunakan IBM SPSS Modeler. *Thesis*. Surabaya, Indonesia: Universitas Kristen Petra.
- Suryawan, I. G. (2022). Developing adaptive ANN models for property valuation. *Journal of Advanced Research in Real Estate*, 12(2), 98-110.
- Tanamal, R., Minoque, N., Wiradinata, T., Soekamto, Y., & Ratih, T. (2023). House price prediction model using random forest in Surabaya city. *TEM Journal*, 12(1), 126-132. <https://doi.org/10.18421/TEM121-17>
- Wang, B. (2023). The Effect of Proximity to Universities on House Prices after the COVID Outbreak. *Journal of Risk and Financial Management*, 16(3), 167. <https://doi.org/10.3390/jrfm16030167>
- Yanindah, A. (2021). An insight into youth unemployment in Indonesia. *The 1st International Conference on Data Science and Official Statistics*, 666-682. Jakarta: Politeknik Statistika STIS.
- Zietz, J., Zietz, E. N., & Sirmans, G. S. (2008). Determinants of house prices: A quantile regression approach. *Journal of Real Estate Finance and Economics*, 37(4), 317-333. <https://doi.org/10.1007/s11146-007-9053-7>