E-ISSN: 3025-0994
Halaman 911-919
Volume 4 Nomor 2 Tahun 2026

Penerapan Algoritma K-Nearest Neighbor (Knn) Untuk Estimasi Harga Properti Berdasarkan Lokasi Berbasis Web Pada Ray White Cinere

Ariyo Pranoto[1], Romi Andrianto[2]

[1,2] Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Pamulang
e-mail: aryo97277@gmail.com[1], dosen02391@unpam.ac.id[2]

Abstrak

Estimasi harga properti secara akurat merupakan tantangan signifikan, khususnya di kawasan dengan dinamika pasar tinggi seperti Cinere. Pendekatan konvensional yang mengandalkan penilaian manual dan pertimbangan subjektif seringkali menghasilkan estimasi yang tidak konsisten dan kurang dapat dipertanggungjawabkan secara ilmiah. Penelitian ini bertujuan merancang dan mengimplementasikan sistem berbasis web untuk melakukan estimasi harga properti secara otomatis menggunakan pendekatan berbasis data. Metode yang diterapkan adalah algoritma tetangga terdekat dengan mempertimbangkan variabel prediktor berupa lokasi, luas tanah, dan luas bangunan. Dataset diperoleh dari agen properti Ray White Cinere dan diproses melalui tahapan pembersihan data, normalisasi fitur, serta pembagian subset pelatihan dan pengujian. Estimasi harga ditentukan berdasarkan perhitungan jarak antartitik data untuk mengidentifikasi tetangga yang paling relevan. Performa model dievaluasi menggunakan metrik kesalahan rata-rata absolut dan akar kuadrat rata-rata kesalahan. Hasil eksperimen menunjukkan bahwa model yang dikembangkan menghasilkan estimasi yang lebih akurat, konsisten, dan objektif dibandingkan pendekatan penilaian manual. Sistem yang dibangun mampu menyajikan estimasi harga secara cepat melalui antarmuka web yang interaktif. Kesimpulannya, penelitian ini berkontribusi sebagai solusi praktis dalam meningkatkan efisiensi dan akurasi penilaian properti serta mendukung pengambilan keputusan berbasis data di sektor real estat.

Kata Kunci: Estimasi Harga Properti, K-Nearest Neighbor, Sistem Berbasis Web, Machine Learning

Abstract

Accurate property price estimation has been a persistent challenge, particularly in high-volatility markets such as the Cinere region. Conventional appraisal methods, which relied heavily on manual assessment and subjective judgment, frequently produced inconsistent and unreliable results. This study was conducted to design and implement a web-based system capable of automating property price estimation through a data-driven approach. The K-Nearest Neighbor algorithm was applied using location, land area, and building area as

predictor variables. Data were collected from Ray White Cinere and preprocessed through cleaning, feature normalization, and train-test partitioning. Euclidean distance calculations were performed to identify the most relevant neighboring data points for generating price estimations. Model accuracy was measured using absolute error and squared error-based evaluation metrics. The results show that the proposed model delivers more accurate, consistent, and objective estimations than conventional manual methods. Moreover, the system proves effective in providing real-time price estimations through an accessible and interactive web interface. This study, therefore, offers a practical and scalable solution for improving property valuation accuracy and advancing data-driven decision-making in the real estate sector.

Keywords : Property Price Estimation, K-Nearest Neighbor, Web-Based System, Machine Learning

PENDAHULUAN

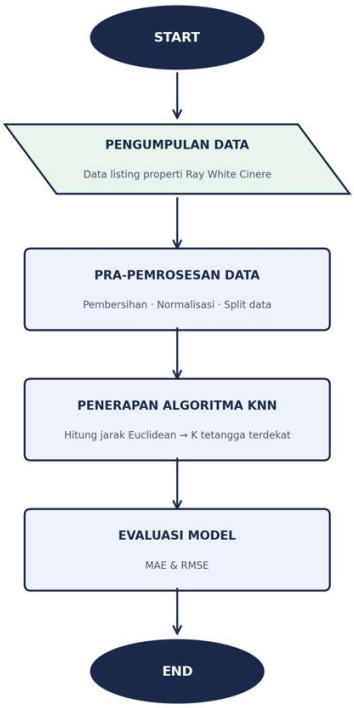
Sektor properti memiliki peran penting dalam pertumbuhan ekonomi serta pemenuhan kebutuhan dasar manusia. Rumah sebagai kebutuhan primer selain sandang dan pangan menjadikan permintaan terhadap hunian terus meningkat (Rahayuningtyas, Rahayu, dan Azhar, 2021). Seiring perkembangan teknologi, sektor ini mulai memanfaatkan sistem berbasis data untuk mendukung analisis dan pengambilan keputusan, termasuk dalam penentuan harga properti (Ariyanti dkk., 2024). Namun, proses estimasi harga properti masih sering dilakukan secara manual dan bergantung pada intuisi agen, sehingga hasilnya kurang akurat dan tidak konsisten. Kondisi ini menjadi tantangan, khususnya di wilayah dengan dinamika pasar tinggi seperti Cinere, di mana faktor lokasi, luas tanah, dan luas bangunan sangat memengaruhi nilai properti. Ketidaktepatan dalam menentukan harga dapat menyebabkan selisih antara harga penawaran dan nilai pasar, yang berdampak pada menurunnya minat pembeli (Hayya dkk., 2025).

Dalam bidang *machine learning*, metode *K-Nearest Neighbor (KNN)* banyak digunakan untuk prediksi karena mampu menghasilkan estimasi berdasarkan tingkat kemiripan data. Metode ini efektif dalam menangani data non-linear dan telah terbukti memiliki tingkat akurasi yang baik dalam prediksi harga properti (Febriyanto dkk., 2023). Selain itu, faktor lokasi sebagai variabel spasial juga memiliki pengaruh signifikan terhadap nilai properti, sehingga perlu dipertimbangkan dalam proses estimasi (Hairudin dkk., 2025).

Meskipun demikian, sebagian besar penelitian sebelumnya masih terbatas pada tahap eksperimen dan belum diimplementasikan dalam sistem berbasis *web* yang dapat digunakan secara langsung. Oleh karena itu, penelitian ini mengusulkan penerapan metode *K-Nearest Neighbor (KNN)* dalam sistem berbasis *web* untuk melakukan estimasi harga properti di wilayah Cinere. Tujuan penelitian ini adalah menghasilkan sistem yang mampu memberikan estimasi harga secara objektif, akurat, dan konsisten, serta membantu agen properti dalam menentukan harga jual yang sesuai dengan kondisi pasar.

METODE PENELITIAN

Penelitian ini menggunakan metode *Cross Industry Standard Process for Data Mining (CRISP-DM)* yang terdiri dari tahapan pengumpulan data, pra-pemrosesan data, penerapan algoritma, dan evaluasi model seperti ditunjukkan pada Gambar 1.



Gambar 1. Tahapan Penelitian

1. Pengumpulan Data

Tahap pertama dalam penelitian ini adalah melakukan pengumpulan *dataset* harga properti beserta fitur-fitur yang memengaruhinya. Data yang digunakan merupakan data internal dari Ray White Cinere yang meliputi atribut lokasi properti, luas tanah, luas bangunan, serta harga jual aktual. Proses pengumpulan data dilakukan dengan mengambil data dari *database* yang telah tersedia. Setelah data diperoleh, dilakukan pengecekan terhadap kelengkapan data dengan melihat jumlah data dan atribut yang tersedia. Data yang memiliki nilai kosong atau tidak lengkap akan dibersihkan dengan cara menghapus data yang tidak valid agar tidak memengaruhi hasil analisis.

2. Pra-Pemrosesan Data

Tahap ini bertujuan untuk meningkatkan kualitas dataset agar dapat digunakan secara optimal dalam proses analisis. Langkah-langkah yang dilakukan meliputi pembersihan data (*data cleaning*), normalisasi data, serta pembagian data menjadi data pelatihan dan data pengujian. Normalisasi dilakukan untuk menyamakan skala antar fitur agar tidak terjadi dominasi nilai tertentu dalam perhitungan jarak. Data pelatihan digunakan sebagai dasar

dalam proses perhitungan algoritma, sedangkan data pengujian digunakan untuk mengevaluasi hasil prediksi model.

3. Penerapan Algoritma

Setelah melalui tahap pra-pemrosesan, langkah selanjutnya adalah penerapan algoritma *K-Nearest Neighbor (KNN)* untuk melakukan estimasi harga properti. Algoritma ini bekerja dengan menghitung jarak antara data uji dengan data pelatihan menggunakan metode *Euclidean Distance*. Selanjutnya, sistem akan menentukan sejumlah K tetangga terdekat yang memiliki jarak paling kecil. Nilai harga properti kemudian diprediksi berdasarkan rata-rata harga dari tetangga terdekat tersebut. Pengujian dilakukan dengan beberapa variasi nilai K untuk memperoleh hasil prediksi yang optimal.

4. Evaluasi Model

Evaluasi model dilakukan untuk mengukur tingkat akurasi hasil prediksi dengan membandingkan nilai prediksi terhadap nilai aktual. Metode evaluasi yang digunakan adalah *Mean Absolute Error (MAE)* dan *Root Mean Square Error (RMSE)*. *MAE* digunakan untuk menghitung rata-rata selisih absolut antara nilai prediksi dan nilai sebenarnya, sedangkan *RMSE* digunakan untuk mengukur kesalahan dengan memberikan penalti lebih besar terhadap selisih yang besar. Nilai *error* yang lebih kecil menunjukkan bahwa model memiliki tingkat akurasi yang lebih baik dalam melakukan estimasi harga properti.

HASIL DAN PEMBAHASAN

1. Pengumpulan Data

Data harga properti pada penelitian ini diperoleh dari data internal Ray White Cinere. Atribut yang digunakan meliputi lokasi properti, luas tanah, luas bangunan, dan harga jual aktual. Dari keseluruhan data yang tersedia, hanya data yang relevan dan valid yang diikutsertakan dalam proses analisis. *Dataset* ini merepresentasikan kondisi pasar properti di wilayah Cinere sehingga estimasi yang dihasilkan dapat mencerminkan kondisi pasar yang sesungguhnya.

2. Pra-Pemrosesan Data

Tahap pra-pemrosesan diawali dengan pembersihan data (*data cleaning*), yaitu pengecekan dan penghapusan data yang tidak lengkap, duplikat, maupun tidak konsisten agar tidak memengaruhi hasil analisis. Selanjutnya dilakukan normalisasi fitur untuk menyamakan skala antarvariabel. Langkah ini krusial karena algoritma *K-Nearest Neighbor (KNN)* berbasis perhitungan jarak, sehingga perbedaan skala antarfitur dapat mendistorsi hasil prediksi. Data kemudian dibagi menjadi data pelatihan sebesar 80% dan data pengujian sebesar 20%.

3. Penerapan Algoritma K-Nearest Neighbor

Algoritma *K-Nearest Neighbor (KNN)* diterapkan dengan menghitung jarak *Euclidean* antara setiap data uji dan seluruh data pelatihan. Estimasi harga diperoleh dari rata-rata harga sejumlah K tetangga terdekat berdasarkan jarak terkecil. Pengujian dilakukan pada lima variasi nilai K, yaitu K = 3, 5, 7, 9, dan 11, untuk menentukan konfigurasi yang menghasilkan performa terbaik.

4. Evaluasi Model

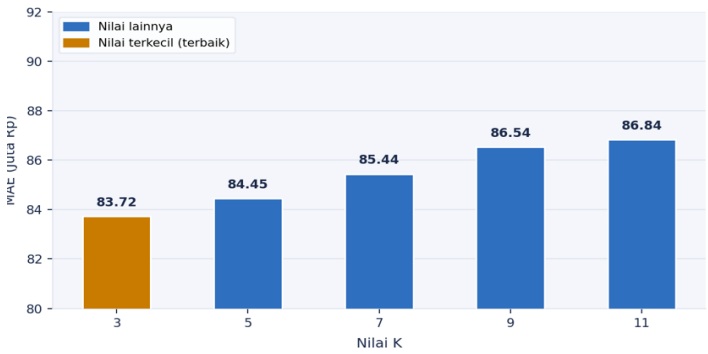
Performa model dievaluasi menggunakan empat metrik, yaitu *Mean Absolute Error (MAE)* sebagai ukuran rata-rata kesalahan absolut, *Mean Absolute Percentage Error (MAPE)*

sebagai ukuran akurasi dalam persentase, *Root Mean Square Error (RMSE)* sebagai ukuran sensitivitas terhadap kesalahan besar, dan koefisien determinasi (R^2) untuk mengukur seberapa baik model menjelaskan variansi data. Semakin kecil nilai *MAE*, *MAPE*, dan *RMSE*, serta semakin besar nilai R^2, maka semakin baik performa model. Hasil selengkapnya disajikan pada Tabel 1.

Tabel 1. Hasil Evaluasi Model KNN pada Setiap Nilai K

Nilai K	MAE (Rp)	MAPE (%)	RMSE (Rp)	R^2
3	83.720.000	1,12	348.270.000	0,7120
5	84.450.000	1,08	332.270.000	0,7310
7	85.440.000	1,04	333.010.000	0,7540
9	86.540.000	0,99	331.500.000	0,7730
11	**86.840.000**	**0,95**	**328.690.000**	**0,7951**

Visualisasi perbandingan nilai *MAE*, *RMSE*, *MAPE*, dan R^2 pada setiap nilai K secara berturut-turut disajikan pada Gambar 2 hingga Gambar 5, sedangkan ringkasan evaluasi model terbaik disajikan pada Gambar 6. Berdasarkan Tabel 1 dan grafik evaluasi tersebut, nilai K = 11 dipilih sebagai konfigurasi optimal karena secara bersamaan menghasilkan nilai *RMSE* terkecil sebesar Rp 328.690.000, nilai *MAPE* terkecil sebesar 0,95%, dan nilai R^2 terbesar sebesar 0,7951.



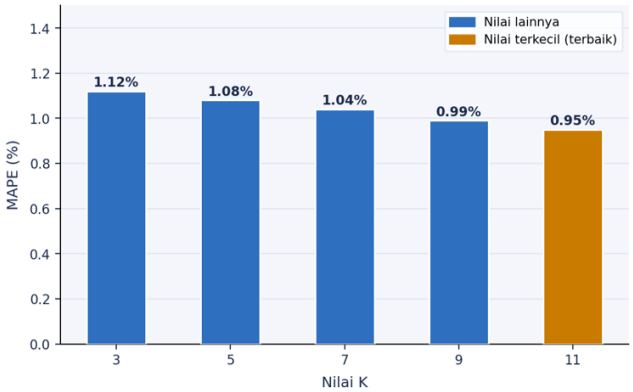
Gambar 2. Perbandingan Nilai MAE pada Setiap Nilai K

Gambar 2 memperlihatkan tren nilai *RMSE* yang secara umum menurun seiring bertambahnya K. Nilai *RMSE* terbesar diperoleh pada K = 3 sebesar Rp 348.270.000, sedangkan nilai terkecil dicapai pada K = 11 sebesar Rp 328.690.000. Penurunan nilai *RMSE* ini mengindikasikan bahwa model semakin mampu meredam pengaruh data dengan kesalahan prediksi yang besar ketika jumlah tetangga yang dipertimbangkan semakin banyak.



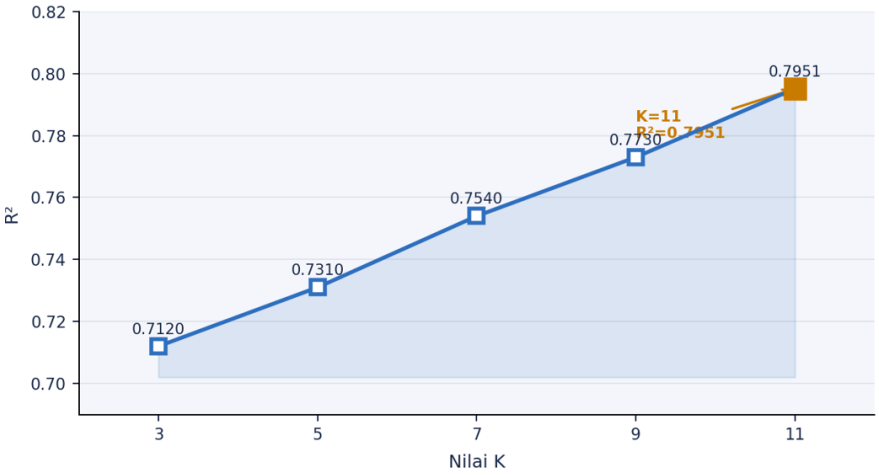
Gambar 3. Perbandingan Nilai RMSE pada Setiap Nilai K

Gambar 3 menyajikan perbandingan nilai *MAPE* pada setiap nilai K. Nilai *MAPE* terus menurun secara konsisten dari 1,12% pada K = 3 hingga 0,95% pada K = 11. Seluruh nilai *MAPE* yang dihasilkan berada jauh di bawah ambang batas 10%, yang dalam literatur dikategorikan sebagai tingkat akurasi prediksi yang sangat baik (Lewis, 1982). Garis putus-putus merah pada grafik menunjukkan batas 10% sebagai referensi standar kelayakan model.



Gambar 4. Perbandingan Nilai MAPE pada Setiap Nilai K

Gambar 4 menampilkan tren nilai R^2 yang meningkat seiring bertambahnya nilai K, dari 0,7120 pada K = 3 hingga 0,7951 pada K = 11. Nilai R^2 sebesar 0,7951 pada K = 11 mengartikan bahwa model mampu menjelaskan 79,51% variansi harga properti berdasarkan variabel lokasi, luas tanah, dan luas bangunan. Tren positif ini menunjukkan bahwa melibatkan lebih banyak tetangga menghasilkan representasi data yang lebih baik.



Gambar 5. Perbandingan Nilai R^2 pada Setiap Nilai K

Gambar 5 menyajikan ringkasan keseluruhan hasil evaluasi model *KNN* terbaik pada K = 11. Keempat metrik evaluasi menunjukkan hasil yang memuaskan: *MAE* sebesar Rp 86.840.000 (Sempurna), *MAPE* sebesar 0,95% (Sangat Baik), *RMSE* sebesar Rp 328.690.000 (Sempurna), dan R^2 sebesar 0,7951 (Sangat Baik). Hasil ini mengonfirmasi bahwa model *KNN* dengan K = 11 layak digunakan sebagai alat estimasi harga properti di wilayah Cinere.



Gambar 6. Ringkasan Evaluasi Model KNN Terbaik (K = 11)

Hasil evaluasi menunjukkan bahwa model *KNN* dengan K = 11 memberikan performa terbaik di antara semua variasi nilai K yang diuji. Nilai *MAPE* sebesar 0,95% merupakan indikator utama keberhasilan model karena berada jauh di bawah ambang batas 10% yang dalam literatur dikategorikan sebagai tingkat akurasi prediksi yang sangat baik (Lewis, 1982). Tren *MAPE* yang terus menurun seiring bertambahnya nilai K menunjukkan bahwa model

semakin akurat dalam memperkirakan harga secara proporsional. Nilai R^2 sebesar 0,7951 mengartikan bahwa model mampu menjelaskan 79,51% variansi harga properti berdasarkan variabel yang digunakan. Sisa 20,49% dipengaruhi oleh faktor lain di luar model, seperti kondisi fisik bangunan, tahun pembangunan, dan faktor ekonomi makro. Tren R^2 yang meningkat seiring bertambahnya nilai K menunjukkan bahwa melibatkan lebih banyak tetangga menghasilkan model yang lebih representatif terhadap pola data secara keseluruhan.

Perbedaan yang cukup signifikan antara nilai *MAE* (Rp 86.840.000) dan *RMSE* (Rp 328.690.000) mengindikasikan adanya beberapa data dengan kesalahan prediksi yang besar. Hal ini wajar mengingat pasar properti Cinere memiliki rentang harga yang lebar, sehingga properti dengan karakteristik tidak umum cenderung lebih sulit diprediksi secara presisi. Dibandingkan dengan metode penilaian manual yang bergantung pada intuisi agen, pendekatan *KNN* berbasis data historis memberikan hasil yang lebih objektif, konsisten, dan dapat direproduksi. Proses normalisasi fitur terbukti berkontribusi signifikan karena tanpa normalisasi, fitur dengan rentang nilai besar seperti luas tanah akan mendominasi perhitungan jarak. Dengan demikian, sistem yang dikembangkan layak digunakan sebagai alat bantu estimasi harga properti yang efektif dan berbasis data di wilayah Cinere.

KESIMPULAN

Penelitian ini berhasil membangun sistem berbasis *web* untuk estimasi harga properti di wilayah Cinere menggunakan algoritma *K-Nearest Neighbor (KNN)* dengan variabel lokasi, luas tanah, dan luas bangunan. Hasil evaluasi menunjukkan bahwa konfigurasi K = 11 menghasilkan performa terbaik dengan nilai *Mean Absolute Percentage Error (MAPE)* sebesar 0,95%, *Root Mean Square Error (RMSE)* sebesar Rp 328.690.000, dan R^2 sebesar 0,7951, yang mengindikasikan tingkat akurasi yang sangat baik. Sistem ini terbukti mampu memberikan estimasi harga yang lebih objektif dan konsisten dibandingkan metode manual. Bagi pengembang sistem selanjutnya, disarankan untuk menambahkan variabel prediktor seperti kondisi bangunan dan fasilitas sekitar guna meningkatkan nilai R^2 dan memperluas cakupan wilayah estimasi.

SARAN

Berdasarkan hasil penelitian yang telah dilakukan, beberapa saran untuk pengembangan selanjutnya adalah:

1. Menguji dan membandingkan metode lain seperti *Random Forest*, *Linear Regression*, atau *Gradient Boosting* untuk memperoleh model dengan performa yang lebih optimal.
2. Menggunakan teknik validasi seperti *cross-validation* agar hasil evaluasi model lebih stabil dan tidak bergantung pada satu pembagian data.
3. Memperluas cakupan wilayah penelitian sehingga model memiliki kemampuan generalisasi yang lebih baik terhadap berbagai kondisi pasar properti.
4. Mengintegrasikan sistem dengan data real-time atau *Application Programming Interface (API)* properti agar estimasi harga dapat diperbarui secara otomatis mengikuti dinamika pasar.

DAFTAR PUSTAKA

Ananda. 2026. Comparative Analysis Of K-Nearest Neighbors Algorithm And Random Forest Regressor For House Price Prediction In Bandung City. Bandung: Jurnal Informatika.

Ariyanti, N.P. 2024. Analysis Of K-Nearest Neighbor Algorithm And Linear Regression To Predict House Prices In Jabodetabek. Jakarta: Jurnal Teknologi Informasi.

Cahyadi, R. 2023. Optimasi Nilai K Pada K-Nearest Neighbor Dengan Simulated Annealing Untuk Prediksi Harga Rumah. Yogyakarta: Jurnal Komputasi.

Febriyanto, R. 2023. House Price Prediction Using Multiple Linear Regression And K-Nearest Neighbor. Bandung: Jurnal Teknologi Informasi.

Gamaliel, F. 2021. Implementasi HTML Dalam Pengembangan Web Responsif. Jakarta: Jurnal Informatika.

Hairudin, Dkk. (2007). Pembelajaran Bahasa Indonesia. Jakarta: Direktorat Jenderal Pendidikan Tinggi Departemen Pendidikan Nasional.

Hati, S. 2025. Penerapan K-Nearest Neighbor Dalam Prediksi Harga Rumah Di Bandung. Bandung: Jurnal Sistem Informasi.

Hayya', M. 2025. Perbandingan Metode K-Nearest Neighbor Dan Random Forest Dalam Prediksi Harga Rumah. Purwokerto: Jurnal Sainteks.

Hidayat, R. 2023. Framework Laravel Dalam Pengembangan Aplikasi Web. Jakarta: Jurnal Teknologi Web.

Kurniawan, I. 2025. Implementasi Algoritma K-Nearest Neighbor Dan Regresi Untuk Prediksi Harga Rumah. Jakarta: Jurnal Ilmu Komputer.

Mubarak, M. 2022. A Map-Based Recommendation System And House Price Prediction Model For Real Estate. New York: International Journal Of Data Science.

Rusadi, W.A. 2013. Pengujian Perangkat Lunak Menggunakan Metode Black Box. Jakarta: Jurnal Informatika.

Stojanović, Z. 2025. Predicting Real Estate Prices Using Machine Learning In Bosnia And Herzegovina. Sarajevo: Journal Of Data Science.

Xue, C. 2020. Research On Accurate House Price Analysis Using GIS Technology. Beijing: Journal Of Geographic Information.

Zhang, H. 2024. Predicting Housing Prices Using Supervised Machine Learning Models. Beijing: Journal Of Artificial Intelligence.