Sainteks
ISSN: 2686-0546
Volume 22 No 1, April 2025
DOI: 10.30595/sainteks.v22i1.25998 (99-108)

Perbandingan Metode *K-Nearest Neighbor* (KNN) dan *Random Forest* dalam Memprediksi Harga Rumah

Comparison of K-Nearest Neighbor and Random Forest Methods in Predicting House Prices

Muthi'ah Hayya'[1*], Fida Maisa Hana[2], Agung Prihandono[3]

[1,2,3]*Program Studi Ilmu Komputer, Fakultas Sains dan Teknologi,*

Universitas Muhammadiyah Kudus

[*]corr-author: 32021110001@std.umku.ac.id

ABSTRAK

Semakin meningkatnya populasi, Rumah menjadi kebutuhan primer bagi manusia. Prediksi harga rumah menjadi aspek penting dalam industri properti, karena harga rumah terus mengalami perubahan yang dipengaruhi oleh berbagai faktor. Diantara faktor-faktor tersebut adalah seperti lokasi, luas tanah, luas bangunan, serta jumlah kamar. Dalam pengaruh faktor tersebut, masyarakat memerlukan sistem prediksi harga rumah sesuai dengan kebutuhan. Adapun tujuan dari penelitian ini adalah melakukan analisis perbandingan dari hasil prediksi harga rumah dengan algoritma *Machine Learning* yaitu *K-Nearest Neighbor* (KNN) dan *Random Forest*. Pengujian dilakukan dengan menentukan nilai k pada metode KNN dan *n-estimators* (nilai pohon) pada metode *Random Forest*. Evaluasi dilihat dari nilai *Mean Absolute Error* (MAE), *Mean Squared Error* (MSE), dan *R-Squared score* (R^2) terkecil dari setiap pengujian metode. Hasil dari penelitian ini Random Forest memiliki hasil akurasi terbaik dengan nilai MAE 54,80, MSE 7004,70, dan R^2 0,55 sedangkan KNN memiliki nilai MAE 87,64 , MSE 13314,95 dan R^2 0,14. Hasil tersebut didapat dari pengujian dataset 2020 baris dan 9 kolom fitur. Hasil dari penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan sistem prediksi harga rumah di masa depan.

Kata-kata kunci: prediksi; *Machine Learning*; *K-Nearest Neighbor*; *Random Forest*; perbandingan.

ABSTRACT

As the population increases, houses become a primary need for humans. House price prediction is an important aspect in the property industry, because house prices continue to change which is influenced by various factors. Among these factors are such as location, land area, building area, and number of rooms. In the influence of these factors, people need a house price prediction system according to their needs. The purpose of this study is to conduct a comparative analysis of the results of house price predictions with Machine Learning algorithms, namely K-Nearest Neighbor (KNN) and Random Forest. Testing is carried out by determining the value of k in the KNN method and n-estimators (tree values) in the Random Forest method. Evaluation is seen from the smallest Mean Absolute Error (MAE), Mean Squared Error (MSE), and R-Squared score (R^2) *from each method test. The results of this study Random Forest has the best accuracy results with MAE values of 54.80, MSE 7004.70, and* R^2 *0.55 while KNN has MAE values of 87.64, MSE 13314.95 and* R^2

Sainteks
ISSN: 2686-0546
Volume 22 No 1, April 2025
DOI: 10.30595/sainteks.v22i1.25998 (99-108)

0.14. These results were obtained from testing the 2020 row dataset and 9 feature columns. The results of this study are expected to contribute to the development of a house price prediction system in the future.

Keywords: prediction; Machine Learning; K-Nearest Neighbor; Random Forest; comparison.

PENDAHULUAN

Rumah merupakan kebutuhan primer bagi manusia, selain sandang dan pangan (Rahayuningtyas, Rahayu and Azhar, 2021). Selain berfungsi sebagai tempat tinggal, rumah juga sering dijadikan sebagai alat investasi karena nilainya yang cenderung meningkat dari waktu ke waktu (Saiful, 2021). Berdasarkan *Maslow's Hierarchy of Needs*, rumah masuk dalam kategori kebutuhan fisiologis atau kebutuhan dasar manusia (Rahayuningtyas, Rahayu and Azhar, 2021). Dalam konteks pertumbuhan populasi yang meningkat, prediksi harga rumah menjadi penting tidak hanya bagi pembeli rumah pertama tetapi juga bagi investor properti yang memerlukan informasi yang akurat mengenai perkiraan harga rumah di masa depan (Fitri, 2023).

Harga rumah dipengaruhi oleh banyak faktor, termasuk lokasi, luas tanah dan bangunan, akses ke fasilitas umum, serta kondisi ekonomi global dan lokal (Ariyani *et al.*, 2022). Faktor-faktor ini menunjukkan bahwa harga rumah memiliki karakteristik dinamis dan kompleks, membuatnya menantang untuk diprediksi dengan akurasi tinggi (Saiful, 2021). Dalam menghadapi tantangan ini, metode pembelajaran mesin *(machine learning*) telah berkembang sebagai solusi yang mampu meningkatkan akurasi dalam memprediksi harga rumah.

Machine learning adalah sistem pembelajaran yang berbasis komputasi dan mengolah data sebagai input. Metode ini digunakan untuk mempermudah proses klasifikasi dan identifikasi (Dewi, 2023). Dalam penelitian ini menggunakan algoritma *K-Nearest Neighbor* (KNN) dan *Random Forest Tree* dimana dua algoritma tersebut merupakan salah satu metode dalam *machine laerning*.

Metode *K-Nearest Neighbor* adalah teknik yang mengklasifikasikan data baru dengan mengelompokkan data tersebut bersama data lama berdasarkan jarak terdekat dengan data baru tersebut (Ghani Muttaqin, Auliasari and Santi Wahyuni, 2020). Penelitian terdahulu menunjukkan bahwa KNN dapat memberikan hasil yang cukup baik dalam memprediksi. Misalnya, penelitian yang dilakukan oleh (Sudriyanto, Syahro and Fitriani, 2023) yang berjudul Perbandingan Performa Model Machine Learning Support Vector Machine, Neural Network, Dan K-Nearest Neighbor Dalam Prediksi Harga Saham, menunjukkan hasil model prediksi *K-Nearest Neighbor* (KNN) memiliki nilai RMSE paling rendah, yaitu 0,037. Hal tersebut menunjukkan bahwa model *K-Nearest Neighbor* (KNN) memiliki nilai akurasi tertinggi dalam dalam memprediksi harga saham. Penelitian lain dengan menggunakan rasio yang sama dalam proses klasifikasi tingkat udara, algoritma *K-Nearest Neighbor* mendapatkan nilai akurasi 94.44% sedangkan algoritma Naive Bayes mendapatkan nilai akurasi 86.11% (Budianita *et al.*, 2024).

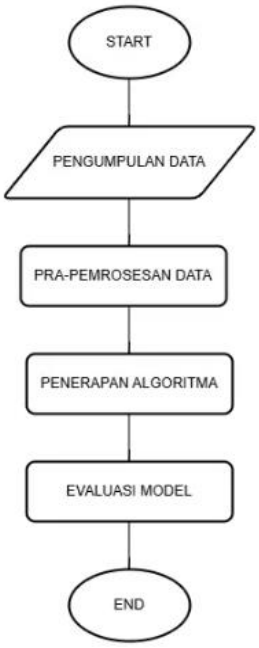
Random Forest adalah sebuah metode yang menggabungkan sejumlah pohon keputusan (decision tree) untuk membentuk suatu kumpulan pohon acak yang digunakan dalam proses klasifikasi atau prediksi data (Suliztia, 2020). Penelitian terdahulu juga menunjukkan bahwa *Random forest* dapat memberikan hasil yang cukup baik dalam memprediksi. Sebagaimana penelitian yang dilakukan oleh (Surya Negara *et al.*, 2023) menunjukkan bahwa model *Random forest* menghasilkan nilai akurasi sebesar 96,38% dibandingkan dengan model yang lain dalam memprediksi harga mobil bekas.

Sainteks
ISSN: 2686-0546
Volume 22 No 1, April 2025
DOI: 10.30595/sainteks.v22i1.25998 (99-108)

Penelitian ini bertujuan untuk melakukan analisis perbandingan antara metode KNN dan Random Forest dalam memprediksi harga rumah, dengan fokus pada akurasi dan performa prediksi dari masing-masing algoritma. Dengan melakukan perbandingan ini, diharapkan dapat ditemukan metode yang lebih tepat dan efektif dalam memodelkan prediksi harga rumah, sehingga bisa digunakan sebagai dasar pertimbangan dalam pengembangan sistem rekomendasi harga properti di masa mendatang.

METODE PENELITIAN

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



Gambar 1. Tahapan penelitian

1. Pengumpulan Data

Tahap pertama dari penelitian ini adalah melakukan pengumpulan dataset harga rumah beserta dengan fitur-fiturnya. Penelitian ini menggunakan dataset yang diambil dari situs *kaggle* berupa harga rumah di wilayah Yogyakarta. Dataset ini terdiri dari 2024 data harga rumah dengan 9 fitur yaitu alamat lokasi, deskripsi rumah, jumlah kamar tidur, jumlah kamar mandi, carport, link website, luas tanah, luas bangunan , dan harga rumah. Proses pengumpulan data dimulai dengan mencari dataset yang sesuai di situs Kaggle, kemudian mengunduhnya dalam format *CSV (Comma-Separated Values)*. Setelah data berhasil diunduh, langkah pertama yang dilakukan adalah pemeriksaan kelengkapan data dengan melihat jumlah entri dan kolom yang tersedia. Data yang tidak memiliki nilai lengkap, seperti harga rumah yang kosong atau informasi fitur yang hilang, akan dibersihkan dengan menghapus atau mengisi nilai yang hilang menggunakan metode tertentu, seperti imputasi rata-rata (*mean imputation*) atau median.

2. Pra Pemrosesan Data

Sainteks
ISSN: 2686-0546
Volume 22 No 1, April 2025
DOI: 10.30595/sainteks.v22i1.25998 (99-108)

Tahap ini bertujuan untuk meningkatkan kualitas dataset agar dapat digunakan secara optimal dalam proses pelatihan model prediksi harga rumah. Langkah-langkah utama dalam pra-pemrosesan data meliputi pembersihan data, penanganan nilai yang hilang, normalisasi fitur, serta transformasi data jika diperlukan.

3. Penerapan Algoritma

Setelah melewati tahap pra-pemrosesan, langkah berikutnya adalah penerapan algoritma K-Nearest Neighbors (KNN) dan Random Forest untuk memprediksi harga rumah berdasarkan fitur yang tersedia. Kedua algoritma ini dipilih karena memiliki pendekatan yang berbeda dalam memodelkan hubungan antara fitur rumah dan harga, sehingga memungkinkan analisis perbandingan yang mendalam terkait keakuratan dan efisiensi masing-masing metode. Algoritma *K-Nearest Neighbors (KNN)* bekerja dengan membandingkan setiap data baru dengan sejumlah *k* data terdekat (*neighbors*) dalam data pelatihan. Nilai harga rumah untuk data baru diprediksi berdasarkan rata-rata harga dari *k* tetangga terdekatnya. *Random Forest* bekerja dengan membuat banyak pohon keputusan dan mengambil rata-rata prediksi dari semua pohon untuk menentukan harga rumah

4. Evaluasi Model

Evaluasi model meggunakan matrix *Mean Absolut Error* (MAE) untuk mengukur rata-rata selisih absolut antara nilai harga rumah yang diprediksi dengan harga sebenarnya. Semakin kecil nilai MAE, semakin akurat model. Lalu *Mean Squared Error* (MSE) untuk menghitung rata-rata dari selisih kuadrat antara harga yang diprediksi dan harga aktual, dan *R-squared* untuk mengukur seberapa baik model dapat menjelaskan variabilitas data. Nilai mendekati 1 menunjukkan bahwa model memiliki kemampuan prediksi yang sangat baik.

HASIL DAN PEMBAHASAN

1. Pengumpulan Data

Pada penlitian ini, data mengenai harga rumah dikumpulkan dari situs *Kaggle*, yang merupakan salah satu *platform* penyedia dataset publik terbesar. Data yang akan digunakan pada penelitian ini adalah data harga rumah di wilayah Yogyakarta tahun 2024 (https://www.kaggle.com/datasets/aliviadyatama19/rumah123-yogya-unfiltered). Data harga rumah yang digunakan terdiri dari 2024 record data dan mencakup 9 variable fitur (Gambar 2).

	price	nav-link	description	listing-location	bed	bath	carport	surface_area	building_area
0	Rp 1,79 Miliar	https://www.rumah123.com/properti/sleman/hos17...	Rumah 2 Lantai Baru di jalan Palagan Sleman Y...	Ngaglik, Sleman	3.0	3.0	2.0	120 m²	110 m²
1	Rp 170 Juta	https://www.rumah123.com/properti/sleman/hos17...	RUMAH BARU DEKAT AL AZHAR DAN UGM	Jombor, Sleman	3.0	2.0	1.0	102 m²	126 m²
2	Rp 695 Juta	https://www.rumah123.com/properti/sleman/hos17...	RUMAH ASRI DAN SEJUK DI BERBAH SLEMAN DEKAT PA...	Berbah, Sleman	2.0	2.0	1.0	100 m²	100 m²
3	Rp 560 Juta	https://www.rumah123.com/properti/sleman/hos17...	Rumah Murah 5 Menit Dari Candi Prambanan Tersi...	Prambanan, Sleman	3.0	1.0	1.0	109 m²	67 m²
4	Rp 200 Juta	https://www.rumah123.com/properti/sleman/hos17...	Rumah Murah Cicilan 1jt Di Moyudan Sleman	Moyudan, Sleman	2.0	1.0	1.0	60 m²	30 m²

Gambar 2. Contoh dataset harga rumah di wilayah Yogyakarta

2. Pra Pemrosesan Data

Sainteks
ISSN: 2686-0546
Volume 22 No 1, April 2025
DOI: 10.30595/sainteks.v22i1.25998 (99-108)

Langkah pertama dalam pra-pemrosesan adalah pembersihan data (*data cleaning*). Pada tahap ini, dataset diperiksa untuk mengidentifikasi entri yang tidak lengkap, data duplikat, atau anomali yang dapat mengganggu proses analisis. Jika ditemukan data yang duplikat, maka data tersebut dihapus untuk menghindari bias dalam model. Selain itu, dilakukan pengecekan terhadap data yang tidak konsisten, misalnya kesalahan dalam pengisian jumlah kamar atau luas bangunan yang tidak masuk akal (Gambar 3).

	price	nav-link	description	listing-location	bed	bath	carport	surface_area	building_area
0	39	1666	725	39	3.0	3.0	2.0	29	11
1	66	1667	579	22	3.0	2.0	1.0	3	20
2	310	1665	572	4	2.0	2.0	1.0	0	1
3	265	1664	1457	50	3.0	1.0	1.0	14	188
4	123	1663	1464	37	2.0	1.0	1.0	271	112

Gambar 3. Hasil proses pembersihan data

Langkah berikutnya adalah penanganan nilai yang hilang (*missing value handling*). Dalam dataset yang diperoleh dari Kaggle, beberapa kolom mungkin memiliki nilai kosong atau tidak tersedia. Jika jumlah data yang hilang kecil, entri tersebut dapat dihapus. Namun, jika data yang hilang cukup banyak, dilakukan imputasi menggunakan metode seperti *mean imputation* (mengisi dengan nilai rata-rata), *median imputation* (mengisi dengan median), atau *mode imputation* (mengisi dengan nilai yang paling sering muncul). Metode ini memastikan bahwa data tetap utuh tanpa kehilangan banyak informasi penting.

Setelah data bersih, dilakukan normalisasi fitur untuk memastikan bahwa semua fitur memiliki skala yang seimbang. Normalisasi sangat penting terutama untuk algoritma yang berbasis jarak seperti *K-Nearest Neighbors* (KNN), karena perbedaan skala antar fitur dapat menyebabkan hasil prediksi yang kurang akurat. Misalnya, luas tanah yang bernilai ribuan meter persegi akan memiliki pengaruh lebih besar dibandingkan jumlah kamar tidur yang hanya bernilai satuan. Oleh karena itu, teknik Min-Max Scaling atau Standard Scaling dari pustaka *scikit-learn* digunakan untuk menyamakan skala fitur. Selain itu, jika terdapat fitur kategorikal seperti lokasi dalam bentuk teks, dilakukan transformasi data dengan mengubahnya menjadi bentuk numerik. Salah satu metode yang umum digunakan adalah *one-hot encoding*, di mana setiap kategori unik akan direpresentasikan dalam bentuk biner (0 dan 1). Transformasi ini diperlukan agar algoritma *machine learning* dapat memahami data dengan lebih baik. Setelah semua proses pra-pemrosesan selesai, dataset yang telah dibersihkan dan dinormalisasi siap untuk digunakan dalam tahap penerapan algoritma KNN dan Random Forest. Dengan pra-pemrosesan yang baik, model yang dihasilkan dapat bekerja lebih optimal dalam memprediksi harga rumah secara akurat. Hasil normalisasi fitur data ditunjukkan pada Gambar 4.

	price	listing-location	bed	bath	carport	surface_area	\
0	39	0.573529	0.041667	0.041667	0.071429	0.084548	
1	66	0.323529	0.041667	0.020833	0.000000	0.008746	
2	310	0.058824	0.020833	0.020833	0.000000	0.000000	
3	265	0.735294	0.041667	0.000000	0.000000	0.040816	
4	123	0.544118	0.020833	0.000000	0.000000	0.790087	

	building_area
0	0.048246
1	0.087719
2	0.004386
3	0.824561
4	0.491228

Gambar 4. Hasil normalisasi fitur data

Sainteks
ISSN: 2686-0546
Volume 22 No 1, April 2025
DOI: 10.30595/sainteks.v22i1.25998 (99-108)

3. Penerapan Algoritma

Algoritma *Machine Learning* yang digunakan dalam penelitian ini adalah *K-Nearest Neighbor* (KNN) dan *Random Forest.*

a. Algoritma *K-Nearest Neighbors (KNN)*

Algoritma *K-Nearest Neighbor* (KNN) bekerja dengan membandingkan setiap data baru dengan sejumlah *k* data terdekat (*neighbors*) dalam data pelatihan. Nilai harga rumah untuk data baru diprediksi berdasarkan rata-rata harga dari *k* tetangga terdekatnya. keakuratan data pelatihan yang bergantung pada perhitungan jarak bisa berkurang karena harus memilih, menguji, dan menetapkan jenis jarak yang digunakan serta atribut yang relevan untuk mendapatkan hasil perhitungan jarak yang optimal. Selain itu, algoritma KNN juga memiliki biaya komputasi yang tinggi, karena memerlukan perhitungan jarak antara setiap instance query dengan seluruh data pelatihan. Kedekatan atau jarak antara data dihitung menggunakan metrik jarak Euclidean yang dirumuskan dalam (1) (Cholil *et al.*, 2021).

$$d = \sqrt{\sum_{i=1}^{p}(x_1 - x_2)^2} \quad (1)$$

Keterangan :
d = jarak
i = variabel data
p = dimensi data
x1 = sampel data
x2 = data uji

b. Algoritma *Random Forest*

Random Forest adalah salah satu teknik yang digunakan dalam klasifikasi dengan membangun sejumlah pohon keputusan. Metode ini dapat meningkatkan akurasi dengan menghasilkan simpul anak untuk setiap node induk melalui proses pemilihan secara acak. Akar pohon dibentuk berdasarkan nilai terkecil dari kriteria pemisahan (*splitting criteria*) pada setiap fitur yang digunakan. Dalam penelitian ini, dilakukan uji coba dengan beberapa nilai *n_estimators* (jumlah pohon) untuk menemukan konfigurasi terbaik. Dalam permasalahan regresi, Random Forest direkomendasikan untuk menggunakan perhitungan *Mean Squared Error* (MSE) dalam menentukan kriteria pemisahan pohon (Lestari and Astuti, 2022) menggunakan (2).

$$MSE_n = \frac{1}{n}\sum_{i=1}^{p}(Y_i - Y_n)^2 \quad (2)$$

Keterangan:
MSE_n= Nilai MSE pada pohon ke-n.
N = Jumlah sampel pada pohon ke-n.
Y_i = Nilai sampel ke-I pada pohon ke-n.
Y_n = Nilai rata-rata sampel pohon ke-n.

4. Evaluasi Model

a. Evaluasi Model KNN

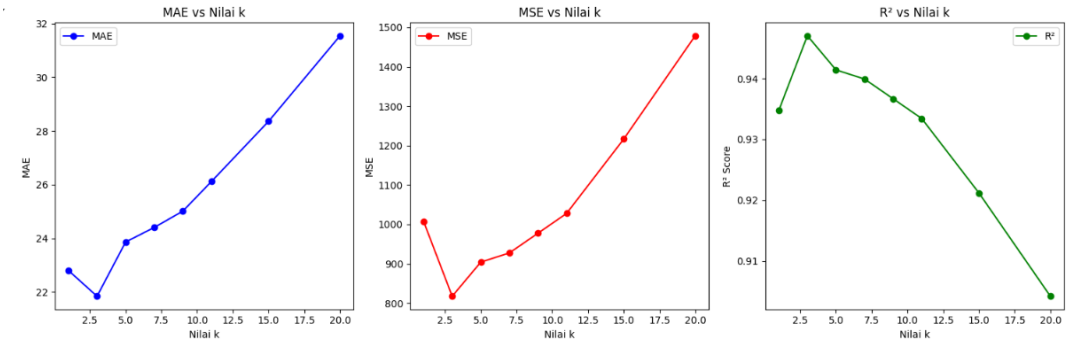
Dalam proses ini, jarak antara data uji dan data latih dihitung menggunakan metrik Euclidean. Selanjutnya, nilai *k* optimal yang telah ditentukan digunakan untuk mengambil rata-rata harga rumah dari tetangga terdekat. Setelah nilai prediksi diperoleh, hasil tersebut dibandingkan dengan harga rumah sebenarnya untuk

Sainteks
ISSN: 2686-0546
Volume 22 No 1, April 2025
DOI: 10.30595/sainteks.v22i1.25998 (99-108)

menghitung metrik evaluasi seperti *Mean Absolute Error* (MAE), *Mean Squared Error* (MSE), dan *R-squared* (R^2). nilai *k* yang diuji disini adalah *k* 1,3,5,7,9,11,15,20. Hasil pengujian akurasi nilai *k* terbaik menunjukkan jika berfokus pada nilai MSE dan R^2 adalah nilai *k* 11 dengan nilai MSE 13314.95 dan R^2 0.14. namun jika berfokus pada nilai MAE, *k* 3 mendapatkan nilai 87.64 meskipun nilai MSE dan R^2 nya kurang baik (Tabel 1) dengan visualisasi grafik seperti pada Gambar 5.

Tabel 1. Hasil Evaluasi Model KNN

Nilai k	MAE (lebih kecil lebih baik)	MSE (lebih kecil lebih baik)	R^2 (lebih besar lebih baik)
1	88.72	19014.88	-0.23 **(sangat buruk)**
3	87.64	14107.01	0.09
5	88.40	13459.25	0.13
7	89.44	13489.66	0.13
9	90.59	13428.71	0.13
11	90.90	13314.95	0.14 **(terbaik dari segi R^2)**
15	93.45	13598.58	0.12
20	94.51	13456.88	0.13



Gambar 5. Grafik hasil evaluasi model KNN

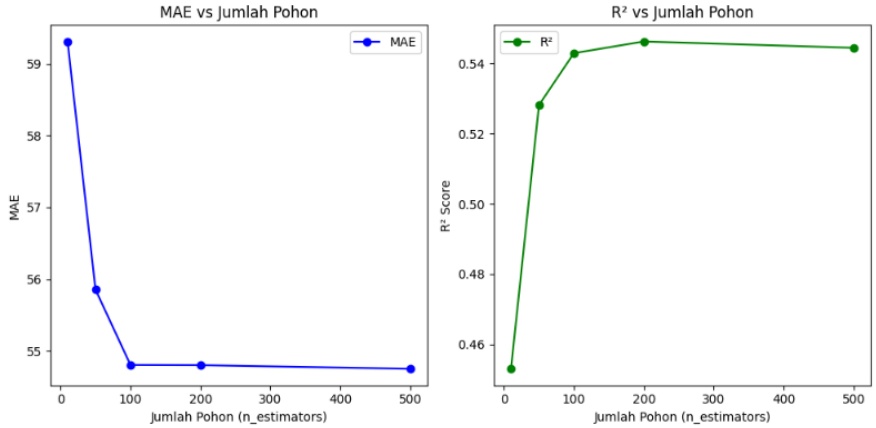
b. Evaluasi Model *Random Forest*

Model dibuat dengan menentukan jumlah *Decision Trees* yang akan digunakan. Jumlah pohon yang lebih besar biasanya meningkatkan akurasi, tetapi juga memperpanjang waktu komputasi. Setiap *Decision Tree* dalam ensemble menghasilkan prediksi masing-masing berdasarkan pola yang telah dipelajari. Prediksi akhir kemudian diperoleh dengan mengambil rata-rata dari seluruh prediksi yang dihasilkan oleh pohon keputusan. Setelah itu, hasil prediksi dibandingkan dengan harga rumah aktual untuk menghitung metrik evaluasi seperti *Mean Absolute Error* (MAE), *Mean Squared Error* (MSE), dan *R-squared* (R^2). Dalam penelitian ini, dilakukan uji coba dengan beberapa nilai *n_estimators* (jumlah pohon) untuk menemukan konfigurasi terbaik. Hasil pengujian akurasi terdapat pada nilai *n_estimators* 200 dengan nilai MAE 54.80, MSE 7004.70, dan R^2 0,55. Hasil tersebut menunjukkan bahwa Model *Random Forest* memiliki akurasi yang lebih baik daripada Model KNN (Tabel 2) dan Gambar 6.

Sainteks
ISSN: 2686-0546
Volume 22 No 1, April 2025
DOI: 10.30595/sainteks.v22i1.25998 (99-108)

Tabel 2. Hasil Evaluasi Model *Random Forest*

n_estimators	MAE (lebih kecil lebih baik)	MSE (lebih kecil lebih baik)	R^2 (lebih besar lebih baik)
10	59.31	8443.06	0.45
50	55.85	7284.06	0.53
100	54.81	7056.03	0.54
200	54.80	7004.70	0.55
500	54.75	7032.52	0.54



Gambar 6. Grafik hasil evaluasi model *Random Forest*

KESIMPULAN

Berdasarkan pengujian model *K-Nearest Neighbor* (KNN) dan *Random Forest* dalam data harga rumah di Yogyakarta, hasil perbandingan prediksi dengan nilai *eror* lebih kecil terdapat pada model *Random Forest.* Hasil evaluasi menunjukkan bahwa Random Forest Regression memiliki akurasi lebih tinggi dibandingkan KNN, dengan nilai MAE terendah sebesar 54.80 dan R^2 tertinggi sebesar 0.55 pada *n_estimators* = 200, sedangkan KNN menunjukkan performa yang kurang optimal dengan MAE hingga 94.51 dan R^2 hanya 0.14 pada k terbaik. Selain itu, penelitian ini juga menekankan pentingnya pra-pemrosesan data, termasuk normalisasi fitur dan penanganan missing values, agar model dapat bekerja lebih optimal. Dengan demikian, dapat disimpulkan bahwa Random Forest Regression lebih direkomendasikan untuk prediksi harga rumah karena memiliki kesalahan yang lebih rendah dan performa yang lebih stabil dibandingkan KNN.

DAFTAR PUSTAKA

Ariyani, V. *et al.* (2022) 'Perbandingan Kinerja Algoritme Naïve Bayes Dan K-Nearest Neighbor (Knn) Untuk Prediksi Harga Rumah', *Jurnal Ilmiah Teknik Elektro*, 24(2), pp. 162–171. Available at: https://ejournal.undip.ac.id/index.php/transmisi.

Budianita, A. *et al.* (2024) 'Komparasi Algoritma K-Nearest Neighbor dan Naive Bayes pada Klasifikasi Tingkat Kualitas Udara Kota Tangerang Selatan', 6(1), pp. 320–327.

Cholil, S.R. *et al.* (2021) 'Implementasi Algoritma Klasifikasi K-Nearest Neighbor (KNN) Untuk Klasifikasi Seleksi Penerima Beasiswa', *IJCIT (Indonesian Journal on*

Computer and Information Technology), 6(2), pp. 118–127. Available at: https://doi.org/10.31294/ijcit.v6i2.10438.

Dewi, L.A. (2023) 'Klasifikasi Machine Learning Untuk Mendeteksi Penyakit Jantung Dengan Algoritma K-Nn, Decision Tree dan Random Forest', *Repository.Uinjkt.Ac.Id* [Preprint]. Available at: https://repository.uinjkt.ac.id/dspace/handle/123456789/71124%0Ahttps://repository.uinjkt.ac.id/dspace/bitstream/123456789/71124/1/LIZKY ASKA DEWI-FST.pdf.

Fitri, E. (2023) 'Analisis Perbandingan Metode Regresi Linier, Random Forest Regression dan Gradient Boosted Trees Regression Method untuk Prediksi Harga Rumah', *Journal of Applied Computer Science and Technology*, 4(1), pp. 58–64. Available at: https://doi.org/10.52158/jacost.v4i1.491.

Ghani Muttaqin, A., Auliasari, K. and Santi Wahyuni, F. (2020) 'Penerapan Metode K-Nearest Neighbor Untuk Prediksi Penjualan Berbasis Web Pada Pt.Wika Industri Energy', *JATI (Jurnal Mahasiswa Teknik Informatika)*, 4(2), pp. 1–6. Available at: https://doi.org/10.36040/jati.v4i2.2728.

Heaton, J. (2015) *Artificial Intelligence for Humans, Volume 3: Neural Networks and Deep Learning*. 1.0. Edited by T. Heaton. Chesterfield, USA: Heaton Research Inc.

Lestari, E.S. and Astuti, I. (2022) 'Penerapan Random Forest Regression Untuk Memprediksi Harga Jual Rumah Dan Cosine Similarity Untuk Rekomendasi Rumah Pada Provinsi Jawa Barat', *Jurnal Ilmiah FIFO*, 14(2), p. 131. Available at: https://doi.org/10.22441/fifo.2022.v14i2.003.

Linguamatics (2019) *What is NLP Text Mining?*, *Linguamatics*.

Mustafidah, H. and Suwarsito, S. (2015) *Model Parameter Jaringan Syaraf Tiruan untuk Pemilihan Algoritma Pelatihan Jaringan Backpropagation yang Paling Optimal*. Purwokerto, Central Java, Indonesia.

Mustafidah, H. and Suwarsito, S. (2016) 'Testing Design of Neural Network Parameters in Optimization Training Algorithm', in I. 2016 (ed.) *International Conference of Result and Community Services, 6th August 2016*. Purwokerto, Indonesia: UMP Press, p. THN. 139-146.

Rahayuningtyas, E.F., Rahayu, F.N. and Azhar, Y. (2021) 'Prediksi Harga Rumah Menggunakan General Regression Neural Network', *Jurnal Informatika*, 8(1), pp. 59–66. Available at: https://doi.org/10.31294/ji.v8i1.9036.

Saiful, A. (2021) 'Prediksi Harga Rumah Menggunakan Web Scrapping dan Machine Learning Dengan Algoritma Linear Regression', *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, (1), pp. 41–50. Available at: https://doi.org/10.35957/jatisi.v8i1.701.

Sofia, M.A., Mustafidah, H. and Suwarsito, S. (2015) 'Basis Data Fuzzy Model Tahani untuk Menentukan Jenis Pakan Ikan Berdasarkan Harga dan Kandungan Gizi Bahan Baku Pakan', *JUITA (Jurnal Informatika)*, III(3), pp. 143–155. Available at: https://doi.org/http://dx.doi.org/10.30595/juita.v3i3.870.

Sudriyanto, S., Syahro, F. and Fitriani, N. (2023) 'Perbandingan Performa Model Machine Learning Support Vector Machine, Neural Network, Dan K-Nearest Neighbors Dalam Prediksi Harga Saham', *Jurnal Advanced Research Informatika*, 2(1), pp. 13–21. Available at: https://doi.org/10.24929/jars.v2i1.2983.

Suliztia, M.L. (2020) 'Penerapan Analisis Random Forest Pada Prototype Sistem Prediksi Harga Kamera Bekas Menggunakan Flask', *Fakultas Matematika Dan Ilmu Pengetahuan Alam*, pp. 1–107.

Surya Negara, E. *et al*. (2023) 'Sulaiman et al, Komparasi Algoritma K-Nearest Neoghbors dan Random Forest …….. 337 Komparasi Algoritma K-Nearest Neighbors dan Random Forest Pada Prediksi Harga Mobil Bekas', *Jurnal JUPITER*, 15(1), pp. 337–346. Available at: www.cardekho.com.