

**ANALISIS KOMPARATIF ALGORITMA K-NEAREST
NEIGHBORS DAN RANDOM FOREST REGRESSOR UNTUK
PREDIKSI HARGA RUMAH DI KOTA BANDUNG**

LAPORAN NON REGULER - SCIENTIST

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana

Program Studi Informatika



Disusun oleh :

DIMAS YUDHISTIRA ANANDA

21.11.4301

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

**ANALISIS KOMPARATIF ALGORITMA K-NEAREST
NEIGHBORS DAN RANDOM FOREST REGRESSOR UNTUK
PREDIKSI HARGA RUMAH DI KOTA BANDUNG**

LAPORAN NON-REGULER - SCIENTIST

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



Disusun oleh :

DIMAS YUDHISTIRA ANANDA

21.11.4301

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

HALAMAN PERSETUJUAN

JALUR NON-REGULER

**ANALISIS KOMPARATIF ALGORITMA K-NEAREST
NEIGHBORS DAN RANDOM FOREST REGRESSOR UNTUK
PREDIKSI HARGA RUMAH DI KOTA BANDUNG**

yang disusun dan diajukan oleh
DIMAS YUDHISTIRA ANANDA
21.11.4301

telah disetujui oleh Dosen Pembimbing
pada tanggal 2, Maret, 2026

Dosen Pembimbing,



FERIAN FAUZI ABDULLOH S.KOM., M.KOM
NIK. 190302276

HALAMAN PENGESAHAN

JALUR NON-REGULER

**ANALISIS KOMPARATIF ALGORITMA K-NEAREST
NEIGHBORS DAN RANDOM FOREST REGRESSOR UNTUK
PREDIKSI HARGA RUMAH DI KOTA BANDUNG**

yang disusun dan diajukan oleh

DIMAS YUDHISTIRA ANANDA

21.11.4301

Telah dipertahankan di depan Dewan Penguji
pada tanggal 2 Maret 2026

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Theopilus Bayu Sasongko, S.Kom., M.Eng
NIK. 190302375

Windha Mega Pradnya Dhuhita, S.Kom., M.KOM
NIK. 190302185

Ferian Fauzi Abdulloh, S.KOM., M.KOM
NIK. 190302276

Laporan ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 2, Maret. 2026

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriani, S.Kom., M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN KARYA

Yang bertandatangan di bawah ini,

Nama mahasiswa : **DIMAS YUDHISTIRA ANANDA**

NIM : **21.11.4301**

Menyatakan bahwa Laporan dengan judul berikut:

**Analisis Komparatif Algoritma K-Nearest Neighbors dan Random Forest Regressor
Untuk Prediksi Harga Rumah Di kota Bandung**

Dosen Pembimbing : **FERIAN FAUZI ABDULLOH S.KOM., M.KOM**

1. Karya tulis ini adalah benar-benar **ASLI** dan **BELUM PERNAH** diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan kegiatan **SAYA** sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab **SAYA**, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini **SAYA** buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidak-benaran dalam pernyataan ini, maka **SAYA** bersedia menerima **SANKSI AKADEMIK** dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 2, Maret, 2026

Yang Menyatakan,



Signature of Dimas Yudhistira Ananda over a stamp.

Dimas Yudhistira Ananda

HALAMAN PERSEMBAHAN

Laporan ini dipersembahkan dengan penuh rasa hormat dan terimakasih kepada kedua penulis Bapak Taswan dan Ibu Mira Widyati yang selalu memberikan doa, dukungan,fasilitas dan motivasi tanpa henti. Ucapan terimakasih juga ddisampaikan kepada dosen pebimbing dan penguji atas bimbingan serta ilmu yang telah diberikan dalam penyusunan laporan ini. Tidak lupa, apresiasi diberikan kepada sahabat dan rekan-rekan yang senantiasa memberikan dukungan dan semangat. Penulis berharap laporan ini dapat memberikan manfaat dan konstibusi yang positif.



KATA PENGANTAR

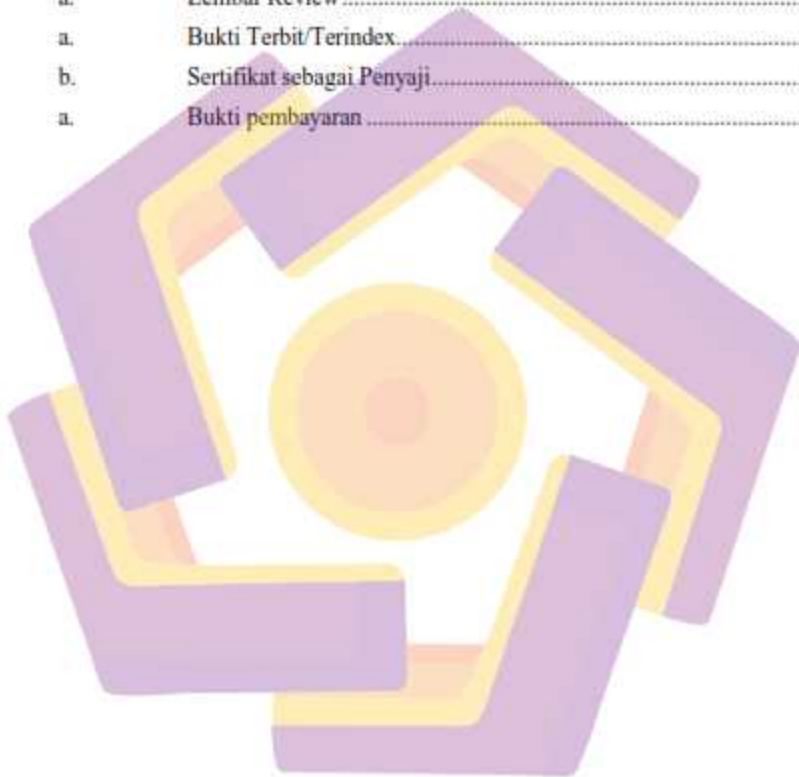
Puji syukur kita panjatkan kehadiran Allah SWT yang telah memberikan rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan laporan skripsi yang berjudul **"Analisis Komparatif Algoritma K-Nearest Neighbors dan Random Forest Regressor Dalam Prediksi Harga Rumah di kota Bandung"** Laporan skripsi ini disusun untuk memenuhi syarat memperoleh gelar Sarjana Komputer Universitas Amikom Yogyakarta. Penulis menyampaikan rasa terima kasih yang tulus serta rasa apresiasi yang mendalam kepada semua pihak yang telah memberikan dukungan dan bantuan yang berharga selama proses penyelesaian skripsi ini. Dengan penuh rasa hormat, ucapan terima kasih yang sebesar-besarnya disampaikan kepada: Prof.

1. . Dr. M. Suyanto., M.M selaku Rektor Universitas Amikom Yogyakarta.
2. Prof. Dr. Kusri S.KOM., M.Kom selaku Dekan Fakultas Ilmu Komputer.
3. Ibu Windha Mega Pradnya Dhuhiya, M.Kom selaku Ketua Program Studi Informatika.
4. Bapak Ferian Fauzi Abdulloh S.KOM., M.Kom selaku Dosen Pembimbing yang telah memberikan ilmu, bimbingan, nasihat, dukungan serta solusi selama proses penyusunan jurnal hingga laporan skripsi ini.
5. Seluruh bapak ibu dosen serta staff karyawan Universitas Amikom Yogyakarta yang memberikan ilmunya selama masa perkuliahan.
6. Kedua orang tua penulis, Ayahanda tercinta Taswan dan Ibunda tercinta Mira widayati yang selalu memberikan cinta, kasih sayang, doa, dukungan, semangat motivasi hingga semua fasilitas keperluan yang diperlukan penulis.
7. Annisa Damayanti, Kevin Armandes dan Mauhiba Salmaa Ghaisani selaku sahabat penulis yang selalu menemani, membantu, mendengarkan keluh kesah penulis, dan memberikan dukungan kepada penulis dari awal perkuliahan hingga penulisan skripsi.

DAFTAR ISI

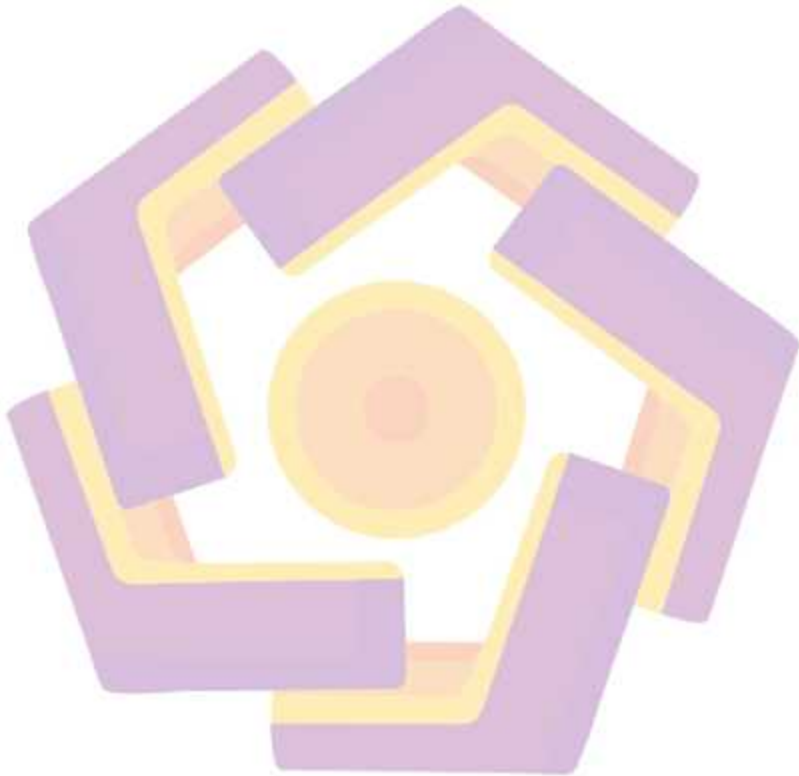
Halaman Judul.....	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN.....	iii
.....	iv
HALAMAN PERSEMBAHAN.....	v
KATA PENGANTAR.....	vi
INTISARI.....	xiv
ABSTRACT.....	xv
BAB I PENDAHULUAN.....	16
1.1. Latar Belakang.....	16
1.2. Rumusan Masalah.....	18
1.3. Batasan Masalah.....	19
1.4. Tujuan.....	19
BAB II TINJAUAN PUSTAKA.....	20
2.1. Studi Mengenai K-Nearest Neighbors (KNN) dalam prediksi harga rumah 20	
2.4 Landasan Teori.....	22
BAB III METODE PENELITIAN.....	25
3.1. Metode.....	25
3.2. Data Collection.....	26
3.3. Analisis data Eksplorasi (EDA).....	28
3.4. Preprocessing Data.....	29
3.4.1 Seleksi dan pembersihan fitur.....	29
.....	29
3.4.2 Encoding Data Kategorikal.....	30
3.4.3 Outlier.....	30
3.4.4 Feature scaling.....	31
3.6. Pemodelan KNN Regressor.....	32
BAB IV KESIMPULAN.....	43
4.1. Kesimpulan.....	43

4.1.	Saran	44
4.2.	Ucapan Terimakasih.....	44
REFERENSI.....		45
CURRICULUM VITAE.....		47
LAMPIRAN DAN BUKTI PENDUKUNG		48
a.	Letter of Acceptance (LOA)	48
a.	Lembar Review	49
a.	Bukti Terbit/Terindex.....	50
b.	Sertifikat sebagai Penyaji.....	50
a.	Bukti pembayaran	51



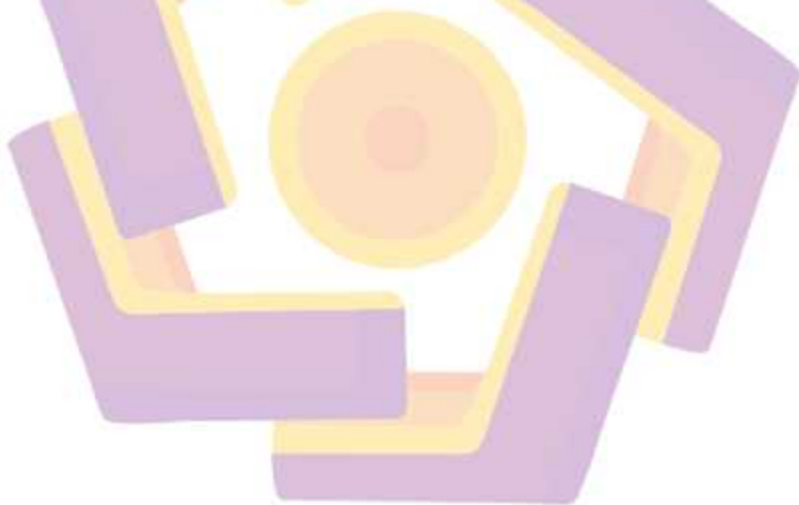
DAFTAR TABEL

Table 1 Identifikasi Dataset	26
Table 2 deskripsi fitur dataset.....	27
Table 3 Hasil dari Grid Search	32
Table 4 Hasil dari pipeline Grid Search	32
Table 5 Ruang Pencarian Hyperparameter	34



DAFTAR GAMBAR

Gambar 1 Flow Chart CRISP-DM	25
Gambar 2 Distribusi Frekuensi Properti Berdasarkan Kecamatan di Kota Bandung	28
Gambar 3 Matriks Korelasi Heatmap	29
Gambar 4 dataframe	29
Gambar 5 Box Plot	31
Gambar 6 Perbandingan MAE Model Random Forest pada Tiga Tahap Pemdoelan	35
Gambar 7 . Perbandingan R^2 Score Model Random Forest pada Tiga Tahap Pemodelan	36
Gambar 8 Tingkat Kepentingan Fitur pada Model Random Forest	37
Gambar 9 Perbandingan R^2 Score Model KNN pada Dua Tahap Pemodelan	38
Gambar 10 Perbandingan R^2 Score KNN pada dua tahap pemodelan	39
Gambar 11 Perbandingan Nilai Aktual vs. Prediksi (Model KNN)	41
Gambar 12 Perbandingan Nilai Aktual vs. Prediksi (Model Random Forest)	41



DAFTAR LAMBANG DAN SINGKATAN

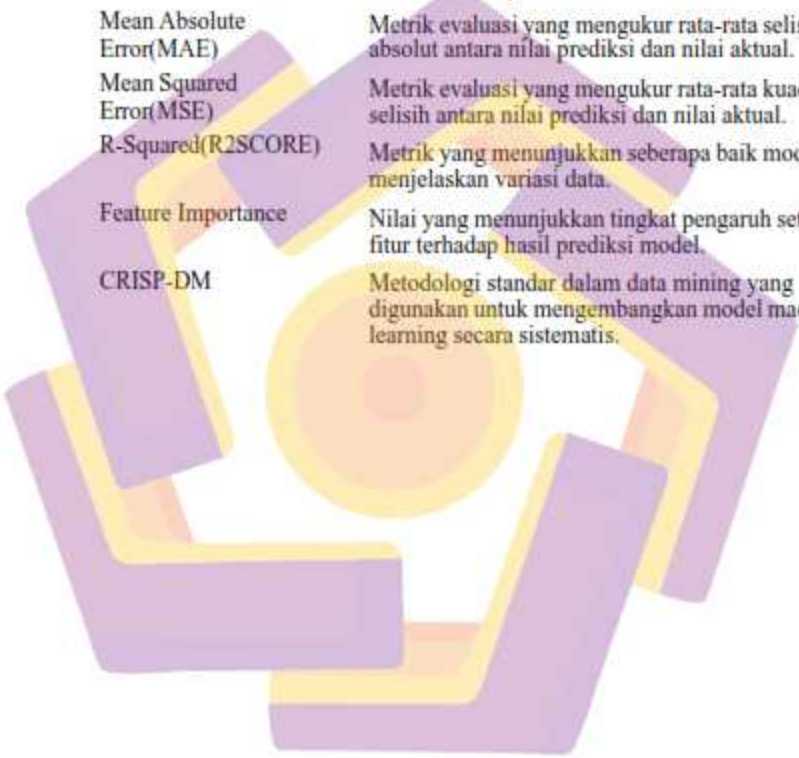


KNN	<i>K-Nearest Neighbors</i>
RF	<i>Random Forest</i>
ML	<i>Machine Learning</i>
EDA	<i>Exploratory Data Analysis</i>
CRISP-DM	<i>Cross-Industry Standard Process for Data Mining</i>
MAE	<i>Mean Absolute Error</i>
MSE	<i>Mean Squared Error</i>
R^2	<i>R-Squared</i>
IQR	<i>Interquartile Range</i>
CV	<i>Cross Validation</i>
$\hat{y}$	<i>Nilai prediksi</i>
Y	<i>Nilai Aktual</i>
K	<i>Jumlah tetangga terdekat pada algoritma KNN</i>
n	<i>Jumlah Data</i>
Σ	<i>Simbol penjumlahan</i>
X	<i>Variabel input atau fitur</i>

DAFTAR ISTILAH



Machine Learning	Teknik untuk mengevaluasi performa model pembelajaran mesin.
Dataset	Kumpulan data yang digunakan sebagai sumber informasi dalam proses pelatihan dan pengujian model.
Regresi	Metode dalam machine learning yang digunakan untuk memprediksi nilai numerik kontinu, seperti harga rumah.
K-Nearest Neighbors (KNN)	Algoritma machine learning berbasis kedekatan yang memprediksi nilai berdasarkan rata-rata nilai dari sejumlah tetangga terdekat.
Random forest	Algoritma ensemble learning yang menggunakan banyak pohon keputusan untuk meningkatkan akurasi dan stabilitas prediksi.
Decision Tree	Model prediksi berbentuk struktur pohon yang digunakan untuk mengambil keputusan berdasarkan fitur-fitur tertentu.
Feature	Variabel atau atribut yang digunakan sebagai input dalam model machine learning.
Target Variabel	Variabel yang menjadi tujuan prediksi model, dalam penelitian ini adalah harga rumah.
Preprocessing Data	Proses persiapan data sebelum digunakan dalam pemodelan, termasuk pembersihan, transformasi, dan normalisasi data.
Outlier	Data yang memiliki nilai jauh berbeda dari data lainnya dan dapat mempengaruhi kinerja model.
One-Hot Encoding	Teknik untuk mengubah data kategorikal menjadi bentuk numerik agar dapat digunakan oleh model machine learning.
Feature Scaling	Proses menyesuaikan skala fitur agar memiliki rentang nilai yang seragam.
Hyperparameter Tuning	Proses mencari kombinasi parameter terbaik untuk meningkatkan performa model.



Grid search	Metode untuk menemukan kombinasi hyperparameter terbaik dengan mencoba semua kemungkinan parameter.
Training Data	Data yang digunakan untuk melatih model agar dapat mengenali pola.
Testing Data	Data yang digunakan untuk mengevaluasi kinerja model setelah pelatihan.
Mean Absolute Error(MAE)	Metrik evaluasi yang mengukur rata-rata selisih absolut antara nilai prediksi dan nilai aktual.
Mean Squared Error(MSE)	Metrik evaluasi yang mengukur rata-rata kuadrat selisih antara nilai prediksi dan nilai aktual.
R-Squared(R2SCORE)	Metrik yang menunjukkan seberapa baik model menjelaskan variasi data.
Feature Importance	Nilai yang menunjukkan tingkat pengaruh setiap fitur terhadap hasil prediksi model.
CRISP-DM	Metodologi standar dalam data mining yang digunakan untuk mengembangkan model machine learning secara sistematis.

INTISARI

Pertumbuhan populasi yang pesat dan ekspansi urban yang berkelanjutan Di Bandung telah berkontribusi terhadap volatilitas dan eskalasi harga perumahan, Menciptakan tantangan signifikan bagi transparansi pasar dan keterjangkauan, Penelitian ini bertujuan untuk mengembangkan dan mengevaluasi model pembelajaran mesin untuk memprediksi harga rumah di wilayah Bandung menggunakan dataset publik yang terdiri dari 7.609 catatan properti mengikuti Metodologi CRISP-DM, penelitian ini mencakup eksplorasi data, , praproses (penanganan outlier menggunakan IQR, one-hot encoding, dan standardisasi fitur), pelatihan model, dan evaluasi kinerja. Dua model regresi—K-Nearest Neighbors (KNN) Regressor dan Random Forest (RF) Regressor—dibandingkan melalui penyetelan hyperparameter sistematis menggunakan teknik Grid Search dan Random Search. Hasil eksperimen menunjukkan bahwa Random Forest Regressor mencapai kinerja terbaik dengan skor R^2 sebesar 0,7838 dan mean absolute error (MAE) sekitar Rp 399,7 juta, mengungguli model KNN yang telah dioptimalkan. Analisis feature importance juga menunjukkan bahwa luas tanah, luas bangunan, dan lokasi adalah prediktor paling berpengaruh terhadap harga properti. Temuan ini menyoroti efektivitas metode ensemble dalam menangani data real-estate yang kompleks dan mendemonstrasikan potensi alat prediktif berbasis pembelajaran mesin untuk mendukung pembeli, penjual, dan pembuat kebijakan dalam membuat keputusan yang terinformasi dan berbasis data di pasar perumahan Bandung.

Kata kunci: Regresor KNN, Regresor Random Forest, Prediksi Harga Rumah, Pembelajaran Mesin, Penilaian Properti, Bandung.

ABSTRACT

The rapid population growth and continuous urban expansion in Bandung have contributed to volatile and escalating housing prices, creating significant challenges for market transparency and affordability. This study aims to develop and evaluate machine-learning models to predict house prices in the Bandung region using a publicly available dataset consisting of 7,609 property records. Following the CRISP-DM methodology, the research includes data exploration, preprocessing (outlier handling using IQR, one-hot encoding, and feature standardization), model training, and performance evaluation. Two regression models K-Nearest Neighbors (KNN) Regressor and Random Forest (RF) Regressor—were compared through systematic hyperparameter tuning using Grid Search and Random Search techniques. The experimental results show that the Random Forest Regressor achieves the best performance with an R^2 score of 0.7838 and a mean absolute error (MAE) of approximately Rp. 399.7 million, outperforming the optimized KNN model. Feature importance analysis also indicates that land area, building area, and location are the most influential predictors of property prices. The findings highlight the effectiveness of ensemble methods in handling complex real-estate data and demonstrate the potential of machine-learning-based predictive tools to support buyers, sellers, and policymakers in making informed and data-driven decisions in the Bandung housing market.

Keyword: KNN Regressor, Random Forest Regressor, House Price Prediction, Machine Learning, Real Estate Valuation, Bandung

