

**OPTIMASI MODEL XGBOOST MENGGUNAKAN GRID SEARCH CROSS
VALIDATION UNTUK PENINGKATAN AKURASI PREDIKSI HARGA
RUMAH**

SKRIPSI

Diajukan Untuk Memenuhi Sebagian Syarat Guna
Memperoleh Gelar Sarjana Komputer (S.Kom.)
Pada Program Studi Sistem Informasi



OLEH :

MUH. JUNAIDI

NPM: 2213030125

FAKULTAS TEKNIK DAN ILMU KOMPUTER
UNIVERSITAS NUSANTARA PERSATUAN GURU REPUBLIK INDONESIA
UN PGRI KEDIRI

2026

Skripsi oleh :

MUHL. JUNAIDI

NPM : 2213030125

Judul

**OPTIMASI MODEL XGBOOST MENGGUNAKAN GRID SEARCH CROSS
VALIDATION UNTUK PENINGKATAN AKURASI PREDIKSI HARGA RUMAH**

Telah disetujui untuk diajukan Kepada
Panitia Ujian/Sidang Skripsi Program Studi Sistem Informasi
FTIK UN PGRI Kediri

Tanggal: 10 Juni 2026

Pembimbing I



Rina Firliana, M.Kom

NIDN. 731087703

Pembimbing II



Arie Nugroho, S.Kom, M.M, M.Kom

NIDN. 712108103

Skripsi oleh :

MUHL. JUNAIDI

NPM : 2213030125

Judul

**OPTIMASI MODEL XGBOOST MENGGUNAKAN GRID SEARCH CROSS
VALIDATION UNTUK PENINGKATAN AKURASI PREDIKSI HARGA RUMAH**

Telah dipertahankan di depan Panitia Ujian/Sidang Skripsi

Program Studi Sistem Informasi FTIK UN PGRI Kediri

Tanggal: 25 Juni 2026

Dan Dinyatakan telah Memenuhi Persyaratan

Panitia Penguji :

- | | | |
|--------------|-----------------------------------|---------|
| 1. Ketua | : Rina Firliana, M.Kom | [.....] |
| 2. Penguji 1 | : Rini Indriati, S.Kom, M.Kom | [.....] |
| 3. Penguji 2 | : Arie Nugroho, S.kom, M.M, M.Kom | [.....] |



Mengetahui,

Dekan Fakultas Teknik dan Ilmu Komputer

Dr. Sulistiono, M.Si

NIDN. 0007076801

PERNYATAAN

Yang bertanda tangan di bawah ini saya,

Nama : Muh. Junaidi
Jenis Kelamin : Laki-laki
Tempat/tgl. lahir : Nganjuk/19 Juni 2002
NPM : 2213030125
Fak/Prodi. : FTIK/ S1-Sistem Informasi

menyatakan dengan sebenarnya, bahwa dalam Skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu perguruan tinggi, dan sepanjang pengetahuan saya tidak terdapat karya tulis atau pendapat yang pernah diterbitkan oleh orang lain, kecuali yang secara sengaja dan tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Kediri, 16 Juni 2026

Yang Menyatakan



MUH. JUNAIDI

NPM: 2213030125

MOTTO

عِشْ مَا شِئْتَ فَإِنَّكَ مَيِّتٌ وَأَحِبِّ مَنْ أَحَبَبْتَ فَإِنَّكَ مُفَارِقُهُ وَاعْمَلْ مَا شِئْتَ فَإِنَّكَ لَأَقْبِيهِ

"Hiduplah sesukamu karena sungguh engkau pasti mati. Cintailah siapa pun yang engkau suka karena sungguh kalian pasti berpisah. Berbuatlah sesukamu karena sungguh engkau pasti menemui (balasan) perbuatanmu itu." [HR al-Baihaqi]

ABSTRAK

Muh. Junaidi: Optimasi Model XGBoost Menggunakan Grid Search Cross Validation Untuk Peningkatan Akurasi Prediksi Harga Rumah, Skripsi, Sistem Informasi, Fakultas Teknik dan Ilmu Komputer UN PGRI Kediri, 2026.

Kata Kunci: *XGBoost, GridSearchCV, Prediksi Harga Rumah, Machine Learning.*

Harga rumah merupakan faktor penting dalam pengambilan keputusan di sektor properti, baik bagi pembeli, penjual, maupun *investor*. Namun, penentuan harga rumah seringkali sulit dilakukan secara akurat karena dipengaruhi oleh berbagai faktor seperti lokasi, luas bangunan, dan jumlah fasilitas. Kompleksitas hubungan antar variabel tersebut menyebabkan metode konvensional kurang efektif, sehingga diperlukan pendekatan berbasis data menggunakan machine learning untuk menghasilkan prediksi yang lebih akurat dan objektif.

Penelitian ini bertujuan untuk membangun model prediksi harga rumah yang akurat serta meningkatkan performa model melalui optimasi hyperparameter. Metode yang digunakan adalah algoritma *XGBoost Regressor* yang dioptimasi menggunakan *Grid Search Cross Validation (GridSearchCV)* dalam sebuah pipeline yang terintegrasi dengan tahap preprocessing. Dataset yang digunakan adalah *Jakarta House Price Dataset*, dengan proses pengolahan meliputi pembersihan data, penanganan *missing value*, transformasi logaritmik pada variabel target, serta *encoding* fitur kategorikal menggunakan *One-Hot Encoding*.

Hasil penelitian menunjukkan bahwa model *XGBoost* yang telah dioptimasi menghasilkan performa yang sangat baik dengan nilai R^2 Score sebesar 0.8889, *Mean Absolute Error (MAE)* sebesar 0.2418, *Mean Squared Error (MSE)* sebesar 0.1300, dan *Root Mean Squared Error (RMSE)* sebesar 0.3606. Selain itu, model ini terbukti lebih unggul dibandingkan model Random Forest sebagai pembanding. Hasil prediksi yang dihasilkan memiliki kedekatan yang tinggi dengan nilai aktual, sehingga menunjukkan bahwa model mampu meminimalkan kesalahan prediksi secara signifikan.

Dengan demikian, penelitian ini menyimpulkan bahwa optimasi *XGBoost* menggunakan *GridSearchCV* mampu meningkatkan akurasi prediksi harga rumah dan memenuhi tujuan penelitian. Namun, penelitian ini masih memiliki keterbatasan pada penggunaan dataset yang hanya mencakup wilayah Jakarta. Oleh karena itu, penelitian selanjutnya disarankan untuk menggunakan dataset yang lebih luas, menambahkan fitur yang lebih kompleks, serta mengeksplorasi algoritma dan metode optimasi lain guna meningkatkan performa model.

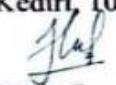
KATA PENGANTAR

Puji Syukur Kami panjatkan Kehadirat Allah Tuhan Yang Maha Kuasa, karena hanya atas perkenan-Nya penyusunan skripsi ini dapat diselesaikan. Skripsi dengan judul “Optimasi Model *XGBoots* Menggunakan *GridSearch Cross Validation* Untuk Peningkatan Akurasi Prediksi Harga Rumah” ini ditulis guna memenuhi sebagian syarat untuk memperoleh gelar Sarjana Komputer, pada Program Studi Sistem Informasi Fakultas Teknik dan Ilmu Komputer UN PGRI Kediri. Pada kesempatan ini diucapkan terima kasih dan penghargaan yang setulus-tulusnya kepada:

1. Rektor Dr. Zainal Afandi, M.Pd. yang senantiasa memberikan dorongan dan fasilitas akademis terbaik kepada seluruh civitas akademika.
2. Dekan Fakultas Dr. Sulistiono, M.Si. atas dukungan dan kebijaksanaannya dalam menjalankan program studi.
3. Ketua Program Studi Dr. Sucipto, M.Kom atas arahan dan fasilitas yang diberikan selama masa perkuliahan dan persiapan penelitian.
4. Ibu Rina Firliana, M.Kom selaku Dosen Pembimbing 1 dan Bapak Arie Nugroho, M.Kom selaku Dosen Pembimbing 2 yang telah meluangkan waktu, memberikan bimbingan, koreksi, dan motivasi tanpa henti, sehingga skripsi ini dapat terselesaikan dengan baik.
5. Orang tua, keluarga, dan seluruh kerabat yang tak pernah lelah memberikan doa, dukungan moril, dan materiil.
6. Teman-teman yang telah memberikan dukungan, motivasi, dan semangat hingga skripsi ini dapat diselesaikan.

Disadari bahwa proposal ini masih memiliki banyak kekurangan dan keterbatasan, baik dari segi penulisan maupun konten. Oleh karena itu, tegur sapa, kritik, dan saran-saran yang membangun dari berbagai pihak sangat diharapkan demi penyempurnaan proposal ini dan pelaksanaan penelitian di masa mendatang.

Kediri, 10 Juni 2026


Muh. Junaidi

NPM: 2213030125

KATA PENGANTAR

DAFTAR ISI

HALAMAN SAMPUL	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN.....	iii
PERNYATAAN.....	iv
MOTTO	v
ABSTRAK	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	viii
DAFTAR GAMBAR.....	xi
DAFTAR TABEL	xii
DAFTAR LAMPIRAN	xiii
BAB I PENDAHULUAN.....	1
A. Latar Belakang	1
B. Batasan Masalah.....	4
C. Rumusan Masalah	4
D. Tujuan Penelitian	4
E. Manfaat Penelitian	5
BAB II KAJIAN TEORI	6
A. Kajian Teori	6
1. Prediksi Harga Rumah dan Machine Learning	6
2. Pra-pemrosesan Data (Data Preprocessing)	7
3. Algoritma Extreme Gradient Boosting (<i>XGBoost</i>)	8
4. Optimasi Hyperparameter Menggunakan <i>Grid Search</i>	9
5. Metrik Evaluasi Model Regresi.....	10
B. Kajian Penelitian Terdahulu.....	11
C. Kerangka Berfikir.....	14

1. Studi Literatur.....	15
2. Identifikasi Kesenjangan dan Arah Solusi	15
3. Tahap Pra-pemrosesan Data (Data Preprocessing)	16
4. Tahap Pemodelan dan Optimasi	16
5. Tahap Evaluasi dan Interpretasi	17
BAB III METODE PENELITIAN	18
A. Alur Penelitian	18
1. <i>Business Understanding</i>	18
2. <i>Data Understanding</i>	19
3. <i>Data Preparation</i>	20
4. <i>Modelling</i>	21
5. <i>Evaluation</i>	21
6. <i>Deployment</i>	22
B. Teknik Pengumpulan Data.....	23
BAB IV HASIL DAN PEMBAHASAN	25
A. Data Hasil Penelitian.....	25
1. Hasil Eksploitasi Data	25
2. Proses transformasi data	28
3. Hasil modeling	30
4. Hasil Perbandingan.....	33
5. Implementasi	34
B. Pembahasan.....	36
1. Model Paling Optimal	36
2. Evaluasi Hasil Modeling	36

3. Kesesuaian Hasil	37
4. Validitas dan Generalisasi Model.....	38
5. Rekomendasi Penerapan Model	38
BAB V.....	40
A. Kesimpulan	40
B. Saran.....	41
DAFTAR PUSTAKA.....	42
LAMPIRAN.....	44

DAFTAR GAMBAR

Gambar 2. 1 Arsitektur <i>XGBoost</i>	8
Gambar 2. 2 <i>Gridsearch CV</i>	9
Gambar 2. 3 Kerangka Berfikir.....	14
Gambar 3. 1 Alur Peneletian.....	18
Gambar 4. 1 Jakarta House Price dataset.....	25
Gambar 4. 2 Struktur Dataset.....	26
Gambar 4. 3 Missing Value Dataset	26
Gambar 4. 4 Statistik Deskriptif Dataset	27
Gambar 4. 5 Distribusi Harga sebelum Transformasi.....	27
Gambar 4. 6 Distribusi Harga Setelah Transformasi	28
Gambar 4. 7 Hasil Penanganan Missing Value.....	28
Gambar 4. 8 Penghapusan Kolom Tidak Relevan	29
Gambar 4. 9 Pipeline.....	29
Gambar 4. 10 Hasil Hyperparameter Tuning.....	30
Gambar 4. 11 Hasil Evaluasi model.....	31
Gambar 4. 12 Grafik <i>Actual vs Predicted</i>	31
Gambar 4. 13 Grafik <i>Residual Plot</i>	32
Gambar 4. 14 Tampilan Awal Aplikasi	34
Gambar 4. 15 Tampilan Hasil Prediksi	35

DAFTAR TABEL

Tabel 3. 1 Tabel Parameter <i>GridSearch CV</i>	21
Tabel 4. 1 Hasil Perbandingan	33

DAFTAR LAMPIRAN

Lampiran 1. Kartu Bimbingan Siakad	44
Lampiran 2. Surat Keterangan Bebas Similarity.....	45
Lampiran 3. Halaman Awal Cek Similarity.....	46
Lampiran 4. Bukti LoA.....	47
Lampiran 5. Berita Acara	48
Lampiran 6. Lembar Revisi.....	49

BAB I

PENDAHULUAN

A. Latar Belakang

Harga rumah merupakan salah satu indikator penting dalam sektor properti yang memengaruhi berbagai keputusan, seperti investasi, pembelian, penjualan, hingga perencanaan pembangunan. Penentuan harga rumah merupakan permasalahan yang kompleks karena melibatkan berbagai faktor non-linear seperti kualitas bangunan, luas area, dan lokasi, sehingga metode konvensional sering kali kurang akurat (Balqis et al., 2026). Dengan adanya prediksi harga yang akurat, berbagai pihak seperti pembeli, penjual, pengembang, dan investor dapat membuat keputusan secara lebih rasional dan terukur. Oleh karena itu, pengembangan sistem prediksi harga rumah yang akurat menjadi kebutuhan yang sangat penting di era data saat ini (Li, 2023). Penerapan *data mining* menjadi fokus utama dalam analisis perbandingan algoritma untuk memprediksi suatu nilai atau kategori berdasarkan fitur-fitur yang tersedia (Kurnia et al., 2024).

Perkembangan teknologi di bidang *machine learning* telah membuka peluang untuk melakukan prediksi harga rumah secara otomatis dengan memanfaatkan data historis dan berbagai atribut properti. Dalam perkembangannya, algoritma berbasis *ensemble learning* seperti *Random Forest* dan *Extreme Gradient Boosting (XGBoost)* telah menjadi standar komputasi yang unggul dalam memprediksi nilai maupun harga sewa aset properti karena fleksibilitasnya dalam memproses fitur non-linear (Yantono et al., 2026). Beberapa algoritma yang umum digunakan dalam penelitian prediksi harga rumah meliputi *Linear Regression*, *K-Nearest Neighbor (KNN)*, *Artificial Neural Network (ANN)*, *Random Forest*, hingga metode *ensemble* lanjutan seperti *XGBoost*. Algoritma tersebut mampu mempelajari hubungan *linear* maupun *non-linear* yang kompleks antara fitur-fitur rumah, seperti lokasi, luas bangunan, jumlah kamar, dan variabel lainnya terhadap harga jual rumah (Hajiyianni & Stylianou, 2025). Oleh karena itu, sangat penting untuk mengembangkan strategi yang tepat untuk melakukan prediksi atau peramalan yang akurat (Fitriono et al., 2024).

Pada penelitian terdahulu yang menjadi jurnal acuan, telah dilakukan perbandingan beberapa algoritma *machine learning* dan *deep learning* untuk memprediksi harga rumah, di antaranya *ANN*, *K-Nearest Neighbor Regression*, *Linear Regression*, dan *Random Forest Regressor*. Selanjutnya, model *Random Forest* dioptimalkan menggunakan metode *Optuna Hyperparameter Tuning* dengan pendekatan *Tree-structured Parzen Estimator (TPE)* untuk meningkatkan akurasi. Hasil penelitian tersebut menunjukkan bahwa *Random Forest* yang telah dioptimalkan menghasilkan performa yang lebih baik dibandingkan model lainnya, dengan peningkatan nilai R^2 dan akurasi yang cukup signifikan (Fadzryan et al., 2025).

Meskipun demikian, hasil prediksi yang dihasilkan model tersebut belum sepenuhnya sempurna. Dalam jurnal tersebut masih ditemukan adanya selisih prediksi yang cukup besar pada beberapa data tertentu. Selisih yang cukup besar ini menunjukkan bahwa, meskipun model telah memiliki tingkat akurasi yang tinggi secara keseluruhan, masih terdapat kelemahan dalam memprediksi beberapa data, khususnya pada rumah dengan karakteristik yang lebih kompleks atau berada pada rentang harga tertentu (Hajiyianni & Stylianou, 2025).

Kondisi tersebut menunjukkan adanya kesenjangan antara harapan dan kenyataan. Secara ideal, model prediksi diharapkan mampu memberikan hasil yang mendekati harga aktual untuk seluruh data. Namun pada praktiknya, masih terdapat kasus *under-prediction* maupun *overprediction* yang cukup signifikan. Hal ini dapat disebabkan oleh beberapa faktor, seperti distribusi data yang tidak merata (*skewed distribution*), keberadaan fitur yang kurang relevan atau belum optimal, serta penggunaan *hyperparameter* model yang belum berada pada konfigurasi terbaik (Sharma et al., 2024). Data yang tidak seimbang (*imbalanced data*) dapat terjadi dan memengaruhi hasil prediksi secara signifikan (Nugroho & Harini, 2024). Oleh karena itu, penggunaan *Grid Search CV* untuk *hyperparameter optimization* merupakan langkah penting yang terbukti dapat meningkatkan akurasi model (Azzaria et al., 2025).

Selain itu, data harga rumah biasanya memiliki kombinasi fitur numerik dan kategorikal. Jika proses pra-pemrosesan tidak dilakukan dengan tepat, seperti penghapusan kolom yang tidak relevan, kesalahan dalam pengolahan data kategorikal, maupun tidak

dilakukannya transformasi data yang diperlukan, maka model akan mengalami kesulitan dalam mempelajari pola sebenarnya yang terdapat di dalam data. Oleh karena itu, penelitian harus melakukan beberapa eksperimen pada pengaturan parameter algoritma untuk mencapai hasil yang lebih optimal (Amarya et al., 2025) dan dibutuhkan suatu pendekatan yang lebih sistematis dalam menyiapkan data dan membangun model prediksi (Hajiyianni & Stylianou, 2025).

Berdasarkan permasalahan tersebut, dalam penelitian ini digunakan pendekatan yang sedikit berbeda dari jurnal acuan. Jika jurnal sebelumnya menggunakan Random Forest yang dioptimalkan dengan *Optuna*, maka dalam penelitian ini digunakan *XGBoost Regressor* yang dikenal memiliki performa tinggi pada tugas regresi dan mampu menangani hubungan non-linear secara lebih efektif. Selain itu, dilakukan transformasi logaritmik pada variabel target (harga) untuk mengatasi kemiringan data, serta diterapkan *ColumnTransformer* dengan *OneHotEncoder* untuk mengolah fitur kategorikal. Selanjutnya, dilakukan proses *hyperparameter tuning* menggunakan *GridSearchCV* untuk mendapatkan kombinasi parameter terbaik.

Dengan pendekatan tersebut, penelitian ini diharapkan mampu meningkatkan akurasi prediksi harga rumah dan mengurangi selisih yang signifikan antara harga aktual dan harga prediksi, sebagaimana yang masih ditemukan pada penelitian sebelumnya. Penelitian ini pada akhirnya diarahkan untuk menghasilkan model prediksi harga rumah yang lebih akurat, stabil, dan dapat diandalkan sebagai alat bantu pengambilan keputusan di bidang properti.

Selain itu, untuk meningkatkan validitas dan relevansi penelitian ini tidak hanya mengadopsi metode yang digunakan, tetapi juga mengimplementasikannya pada dataset “*Jakarta House Price Dataset*” sebagai representasi kondisi pasar properti di Indonesia. Langkah ini bertujuan untuk membandingkan performa metode yang diusulkan dengan metode sebelumnya, sekaligus menguji kemampuan generalisasi model dalam konteks data properti di Indonesia. Dengan demikian, penelitian ini tidak hanya berfokus pada peningkatan performa model, tetapi juga pada evaluasi efektivitas metode dalam menghadapi karakteristik data properti yang berbeda dari penelitian sebelumnya.

B. Batasan Masalah

Agar penelitian ini lebih terarah dan tidak meluas, maka diberikan batasan masalah sebagai berikut:

1. Penelitian ini difokuskan pada upaya meningkatkan akurasi prediksi harga rumah dan meminimalkan selisih antara harga aktual dan harga hasil prediksi melalui optimasi model pembelajaran mesin.
2. Objek dalam penelitian ini adalah data harga rumah beserta variabel-variabel yang memengaruhinya, yang digunakan sebagai dasar dalam proses pelatihan dan pengujian model prediksi.
3. Model yang digunakan dalam penelitian ini dibatasi pada algoritma *XGBoost* Regressor yang diintegrasikan dalam *pipeline* dengan tahap *preprocessing* berupa transformasi logaritmik pada variabel target dan *One Hot Encoding* menggunakan *ColumnTransformer*, serta dilakukan optimasi menggunakan *GridSearchCV*.
4. Dataset yang digunakan dalam penelitian ini adalah “*Jakarta House Price Dataset*”. Dataset “*House Price 2023 Dataset*” hanya digunakan sebagai referensi dalam studi literatur dan bukan sebagai data eksperimen..

C. Rumusan Masalah

Berdasarkan identifikasi dan batasan masalah, maka rumusan masalah pada penelitian ini adalah “Bagaimana optimasi model *XGBoost* menggunakan *Grid Search Cross Validation* dalam meningkatkan akurasi prediksi harga rumah?”

D. Tujuan Penelitian

Tujuan penelitian ini adalah untuk mengoptimalkan model *XGBoost* menggunakan metode *Grid Search Cross Validation* dalam meningkatkan akurasi prediksi harga rumah. Optimasi dilakukan melalui pencarian kombinasi *hyperparameter* terbaik sehingga model mampu menghasilkan performa yang lebih optimal. Keberhasilan model dievaluasi menggunakan metrik R^2 , *Mean Absolute Error (MAE)*, *Mean Squared Error (MSE)*, dan *Root Mean Squared Error (RMSE)* untuk mengetahui kemampuan model dalam menghasilkan prediksi yang mendekati nilai harga rumah sebenarnya.

E. Manfaat Penelitian

Berdasarkan uraian diatas berikut adalah manfaat dari penelitian ini:

1. Manfaat Teoritis

Penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan ilmu pengetahuan, khususnya di bidang *machine learning* dan data mining, melalui penerapan optimasi model *XGBoost* menggunakan *Grid Search Cross Validation* pada kasus prediksi harga rumah. Selain itu, penelitian ini dapat menjadi referensi bagi penelitian selanjutnya yang berkaitan dengan pengembangan model prediksi berbasis *machine learning*.

2. Manfaat Praktis

- a. Bagi masyarakat dan pelaku properti, penelitian ini dapat menjadi alternatif dalam memperkirakan harga rumah secara lebih objektif dan akurat sehingga dapat mendukung proses pengambilan keputusan dalam kegiatan jual beli maupun investasi properti.
- b. Bagi peneliti, penelitian ini dapat menjadi sarana untuk mengembangkan kemampuan dalam menerapkan algoritma *XGBoost*, teknik optimasi *Grid Search Cross Validation*, serta mengevaluasi kinerja model prediksi menggunakan metrik regresi.
- c. Bagi akademisi, hasil penelitian ini diharapkan dapat menjadi bahan referensi dalam pembelajaran maupun penelitian lanjutan yang berkaitan dengan prediksi harga rumah dan penerapan algoritma *machine learning*.

DAFTAR PUSTAKA

- Abdul-rahman, S., & Mutalib, S. (2021). *Advanced Machine Learning Algorithms for House Price Prediction : Case Study in Kuala Lumpur*. 12(12), 736–745.
- Amarya, T. K., Ristyawan, A., & Firliana, R. (2025). *Jurnal Informatika : Jurnal pengembangan IT Optimasi Random Forest untuk Deteksi Dini Penyakit Stroke dengan Data Rekam Medis*. 10(3), 832–840.
<https://doi.org/10.30591/jpit.v10i3.8424>
- Aulia, H., Syifa, N., Nugroho, A., & Firliana, R. (2023). *Perbandingan Algoritma Naïve Bayes Classifier Dan K-Nearest Neighbors Untuk Analisis Sentimen Covid-19 Di Twitter*. 6.
- Azzaria, C., Daniati, E., & Ristyawan, A. (2025). *Peningkatan Akurasi Deteksi Liver Disease melalui Hyperparameter Tuning pada Algoritma Random Forest*. 4(2), 139–147.
- Balqis, F. H., Aini, Q., Studi, P., Informasi, S., Islam, U., Syarif, N., Jakarta, H., & Selatan, K. T. (2026). *Prediksi Harga Rumah menggunakan Algoritma Random Forest Regression House Price Prediction using the Random Forest Regression Algorithm*. 15, 433–445.
- Bhakti, I. N., Sholikhin, A. Z., Abi, M., Daniati, E., & Ristyawan, A. (2024). *Klasifikasi Kategori Berita Menggunakan Naive Bayes*. 1155–1164.
- Cui, J. (2025). *AI in Real Estate : Forecasting House Prices with Advanced Machine Learning Models*. *Daml* 2024, 144–148.
<https://doi.org/10.5220/0013511000004619>
- Dhiwa Aqsha. (2025). *Perbandingan Kinerja Algoritma Extreme Gradient Boosting Dan Random Forest Untuk Prediksi Harga Rumah Di Jabodetabek*. *Jurnal Ilmu Komputer Dan Sistem Informasi*, 13(1), 1–7.
<https://doi.org/10.24912/jiksi.v13i1.32863>
- Dwi, E., Aini, N., Khasanah, R. A., Ristyawan, A., & Neighbors, K. (2024). *Penggunaan Data Mining untuk Prediksi tingkat Obesitas di Meksiko Menggunakan Metode Random*. 8, 1256–1265.
- Fadzryan, R. M., Laksana, E. A., Informatika, T., & Widyatama, U. (2025). *Optimasi Prediksi Harga Rumah dengan Random Forest dan Optuna Hyperparameter Tuning Implementation of Random Forest and Optuna Hyperparameter Tuning Combination to Improve House Price Prediction Accuracy*. 5(6), 1663–1671.
- Fitriyono, D., Wardani, S. A., Al, M. N. B., Ristyawan, A., & Daniati, E. (2024). *Perbandingan Metode Algoritma Decission Tree dan K-Nearest Neighbors untuk Memprediksi Kualitas Air yang dapat dikonsumsi*. 8, 475–484.
- Hajiyianni, N., & Stylianou, V. (2025). *Machine Learning for Real Estate Price Prediction*. 1, 1–11.
- Islami, B. M., & Nugroho, A. (2025). 3) 1, 2, 3. 15(2), 38–50.
- Joseph, V. R., Vakayil, A., Joseph, V. R., & Vakayil, A. (2022). *SPlit : An Optimal Method for Data Splitting SPlit: An Optimal Method for Data Splitting* ABSTRACT. *Technometrics*, 0(0), 1–23.
<https://doi.org/10.1080/00401706.2021.1921037>
- Kurnia, O. D., Fena, E. T. A., Yuliana, D., Ningrum, A., Daniati, E., & Ristyawan, A.

- (2024). *Analisis Perbandingan Algoritma Naïve Bayes dengan K-Nearest Neighbor (KNN) Pada Dataset Mobile Price Classification*. 8, 1174–1183.
- Kurniawan, W., & Indahyanti, U. (2024). *Prediksi Angka Harapan Hidup Penduduk Menggunakan Metode XGBoost*. 2, 1–18.
- Li, H. (2023). *House Price Prediction and Analysis Based on Random Forest and XGBoost Models*. 21, 934–938.
- Nugroho, A., & Harini, D. (2024). *Teknik Random Forest untuk Meningkatkan Akurasi Data Tidak Seimbang*. 2(2), 128–140.
- Rais, Z., S, M. F., & Utomo, A. T. (2025). *Implementation of Machine Learning Algorithm with Extreme Gradient Boosting (XGBoost) Method in Hypertension Level Classification*. 7(1).
- Ristyawan, A., Nugroho, A., & Amarya, T. K. (2025). *Optimasi Preprocessing Model Random Forest Untuk Prediksi Stroke*. 12(1).
- Sharma, H., Harsora, H., & Ogunleye, B. (2024). *An Optimal House Price Prediction Algorithm : XGBoost*. 30–45.
- Syafei, R. M., & Efrilianda, D. A. (2023). *Machine Learning Model Using Extreme Gradient Boosting (XGBoost) Feature Importance and Light Gradient Boosting Machine (LightGBM) to Improve Accurate Prediction of Bankruptcy*. *Recursive Journal of Informatics*, 1(2), 64–72. <https://doi.org/10.15294/rji.v1i2.71229>
- Tr, E., Ristyawan, A., & Nugroho, A. (2025). *Pengaruh Normalisasi Fitur Pada Perbandingan Algoritma Random Forest Regressor dan KNN Regressor dalam Menentukan Harga Smartphone Effect of Feature Normalization on Comparison of Random Forest Regressor and KNN Regressor Algorithms in Determining Smartphone Prices*. 5, 74–83.
- Universitas, T., Luhur, B., Raya, J. C., & Selatan, J. (2024). *XgBoost Hyper-Parameter Tuning Using Particle Swarm Optimization for Stock Price Forecasting*. 9(4), 1179–1195. <https://doi.org/10.26555/jiteki.v9i4.27712>
- Utami, P., Jundi, M., & Sinaga, L. (2025). *Property Price Prediction Using the Random Forest Regression Algorithm*. 22(2), 208–215.
- Utomo, A. D., Hayadi, B. H., & Priyanto, E. (2024). *Enhancing Housing Price Prediction Accuracy Using Decision Tree Regression with Multivariate Real Estate Attributes*. 7(4), 179–188.
- Waciko, K. J., Susanti, L. A., & Nur, R. (2025). *Forecasting Tourist Arrivals in Bali : A Grid Search-Tuned Comparative Study of Random Forest , XGBoost , and a Hybrid RF-XGBoost Model*. 8(3), 251–261.
- Wang, Y., & Zhao, Q. (2022). *House Price Prediction Based on Machine Learning : A Case of King County*. 211(Icfied), 1547–1555.
- Wijaya, H., Hostiadi, D. P., & Triandini, E. (2024). *Optimization XGBoost Algorithm Using Parameter Tuning in Retail Sales Prediction*. *Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI)*, 13(3), 769–786. <https://doi.org/10.23887/janapati.v13i3.82214>
- Yantono, D. J., Puspaningrum, N. R., & Zein, S. A. (2026). *Comparative Analysis of Machine Learning Algorithms for Predicting Commercial Property Rental Prices Analisis Komparasi Algoritma Machine Learning untuk Prediksi Harga Sewa Properti Komersial*. 6(January), 149–155.