

Implementasi Model XGBRegressor Dalam Prediksi Harga Rumah Sebagai Dasar Perencanaan Keuangan Bagi Pasangan Muda

Fajar Sarasati¹, Widi Astuti²

Universitas Nusa Mandiri^{1,2}

fajar.fss@nusamandiri.ac.id¹, widiastuti.wtu@nusamandiri.ac.id²

Diterima (12-09-2025)	Direvisi (01-10-2025)	Disetujui (20-10-2025)
--------------------------	--------------------------	---------------------------

Abstrak - Isu kepemilikan rumah menjadi tantangan utama bagi pasangan muda, terutama karena kenaikan harga rumah yang tidak sebanding dengan pendapatan. Penelitian ini bertujuan untuk memberikan alat bantu dalam pengambilan keputusan finansial bagi pasangan muda yang ingin memiliki rumah pribadi dengan memanfaatkan teknologi machine learning dengan melakukan prediksi harga rumah dengan memanfaatkan algoritma XGBRegressor sebuah metode machine learning berbasis ensemble dengan kemampuan generalisasi yang baik terhadap data kompleks. Dataset yang digunakan dalam penelitian ini meliputi berbagai variabel seperti luas bangunan, luas tanah, jumlah kamar tidur, jumlah kamar mandi dan kapasitas mobil dalam garasi yang memengaruhi harga rumah. Proses penelitian ini meliputi pengumpulan data, pre-processing, pelatihan model, serta evaluasi model dengan menggunakan metrik MAE, MSE dan R² Score. Dimana hasil evaluasi menunjukkan bahwa model XGBRegressor mampu memprediksi harga rumah dengan tingkat akurasi yang cukup baik dengan nilai R² Score 0.77 serta kesalahan prediksi yang ditunjukkan pada nilai MAE dan MSE yang rendah. Penelitian ini mampu memberikan kontribusi bagi pasangan muda yang ingin memiliki rumah pribadi untuk perencanaan finansial masa depan dalam bentuk alat bantu pengambilan keputusan berbasis data secara lebih rasional.

Kata Kunci: Prediksi Harga Rumah, XGBRegressor, Machine Learning, Pasangan Muda

Abstract - Homeownership has become one of the primary needs for young couples, yet it remains difficult to achieve—mainly due to the rising cost of housing each year, which is not in line with the average income of most people. This study aims to provide a decision-support tool to assist young couples in making informed financial plans for owning a private home by utilizing machine learning technology, specifically the XGBRegressor algorithm. XGBRegressor is an ensemble-based machine learning method known for its strong generalization capability on complex datasets. The dataset used in this research includes several variables that influence house prices, such as building area, land area, number of bedrooms, number of bathrooms, and garage capacity. The research process consists of data collection, preprocessing, model training, and model evaluation using MAE, MSE, and R² score as performance metrics. The evaluation results indicate that the XGBRegressor model is capable of predicting house prices with a fairly high level of accuracy, achieving an R² score of 0.77, along with low MAE and MSE values. This study contributes by providing a data-driven decision-support tool to help young couples plan their financial future more rationally when considering purchasing a home.

Keywords: House Price Prediction, XGBRegressor, Machine Learning, Young Couples

I. PENDAHULUAN

Tempat tinggal berupa rumah menjadi kebutuhan dasar manusia. Sebagai makhluk hidup, setiap manusia memiliki kebutuhan yang harus dipenuhi untuk menghidupi 3 macam kebutuhan yakni kebutuhan primer, sekunder dan tersier (Rustini et al., 2025). Kebutuhan primer adalah kebutuhan yang mutlak harus dipenuhi, diantaranya adalah kebutuhan pangan, sandang dan papan (Indonesia, 2024).

Rumah atau papan memiliki peranan penting dalam kehidupan manusia, dengan terpenuhinya kebutuhan dasar manusia berupa rumah maka diharapkan dapat tercapai ketahanan keluarga dan terpenuhinya keamanan dan kenyamanan anggota keluarga (Hardjanto, 2025). Permasalahan kepemilikan rumah telah menjadi isu krusial, terutama bagi generasi muda khususnya pasangan muda yang baru memulai kehidupan berkeluarga.

Harga rumah yang melonjak setiap tahunnya tidak sebanding dengan pendapatan masyarakat membuat impian memiliki rumah pribadi semakin sulit terealisasi.

Harga rumah di Indonesia yang sangat bervariasi tergantung pada lokasi, tipe dan kondisi pasar, namun secara umum harga rumah di Indonesia terus meningkat, terutama di daerah perkotaan dan pusat pertumbuhan ekonomi. Dimana rata-rata kenaikan harga rumah mencapai 5% hingga 10% per tahun (Score, 2024). Pendapatan masyarakat menjadi faktor utama yang memengaruhi daya beli terhadap produk dan jasa termasuk rumah. Semakin besar pendapatan masyarakat maka semakin mampu masyarakat mewujudkan impian memiliki rumah (Roeskamto & Atmadja, 2022). Berdasarkan penelitian yang dilakukan oleh (Roeskamto & Atmadja, 2022) dijelaskan ada tiga faktor yang memengaruhi pergerakan harga rumah yakni (ekonomi, kelembagaan dan demografis). diantara ketiga faktor tersebut faktor ekonomi menjadi faktor penentu pembelian rumah. Pendapatan yang tidak sebanding dengan harga rumah yang semakin melonjak setiap tahunnya menjadikan pasangan muda yang ingin memiliki rumah pribadi harus pandai mengatur rencana finansial untuk kedepannya.

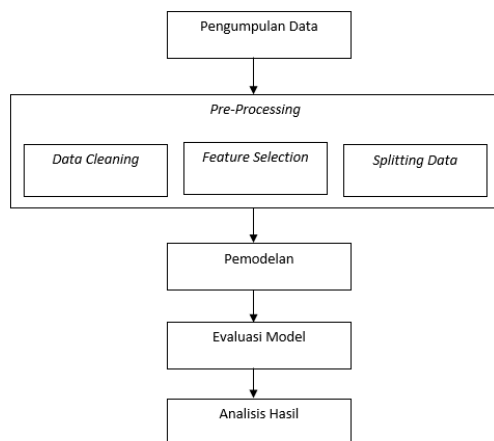
Permasalahan ini mendorong perlunya inovasi berbasis teknologi yang menyediakan informasi yang akurat mengenai prediksi harga rumah untuk membantu masyarakat khususnya pasangan muda dalam membuat keputusan finansial yang terencana. Perkembangan teknologi kecerdasan buatan dan machine learning dapat dimanfaatkan untuk membantu pasangan muda dalam melakukan prediksi terhadap harga rumah di masa mendatang. Salah satu algoritma yang unggul menangani masalah prediksi harga adalah XGBRegressor (Extreme Gradient Boosting Regressore). XGBRegressor ini dipilih karena kemampuannya menangani dataset yang besar dan dikenal karena kinerjanya yang cepat (Al Khairi et al., 2025). Berdasarkan penelitian yang dilakukan (Nuris, 2024) hasil analisis harga rumah dengan memanfaatkan algoritma Regresi Linear menunjukkan tingkat akurasi sebesar 86,41%. Penelitian lainnya dengan tema yang sama juga dilakukan oleh (Wicaksana et al., 2024) yang melakukan prediksi harga rumah berdasarkan luas area tanah menggunakan dataset Kaggle sehingga menghasilkan tingkat akurasi sebesar 65%. Sedangkan pada penelitian lain analisis prediksi harga rumah sesuai spesifikasi harga, luas tanah, luas bangunan, kamar tidur, garasi

dan kota menggunakan Linear Regression menghasilkan tingkat akurasi sebesar 66% (Mu'tashim et al., 2021). Penelitian serupa juga dilakukan oleh (Warjiyono et al., 2024) yang menganalisa prediksi harga jual rumah menggunakan algoritma Random Forest Machine Learning yang menunjukkan nilai akurasi 75,10%. Prediksi harga rumah di Jakarta pusat juga diteliti oleh (Qolbi et al., 2025) dengan menggunakan algoritma Regresi Linear, support vector machine (SVM) dan Random Forest berdasarkan spesifikasi luas tanah dan luas bangunan menghasilkan performa terbaiknya dengan nilai MAE 3023,501.

Sehingga berdasarkan penelitian yang telah dilakukan peneliti sebelumnya, peneliti dapat menguraikan bahwa gap analysis dari penelitian ini adalah jumlah variabel pada dataset yang diteliti lebih banyak dari pada variabel pada penelitian sebelumnya, yakni terdiri dari 7 variabel spesifikasi rumah; nama rumah, harga, luas tanah, luas bangunan, jumlah kamar tidur, jumlah kamar mandi, dan garasi. Selain itu berdasarkan hasil prediksi dengan menggunakan algoritma XGBRegressor menghasilkan performa terbaik dengan nilai R2 Score tinggi dan nilai MAE dan MSE yang kecil. Selain itu pada penelitian ini juga dapat ditampilkan harga rumah berdasarkan spesifikasi rumah yang akan diinginkan oleh pembaca berdasarkan variabel yang sudah dijelaskan. Dimana penelitian ini bertujuan untuk memberikan alat bantu dalam pengambilan keputusan finansial bagi pasangan muda yang ingin memiliki rumah pribadi dengan memanfaatkan teknologi machine learning dengan melakukan prediksi harga rumah. Hasil yang diperoleh dari penelitian ini tidak hanya menunjukkan kemampuan teknik model dalam melakukan prediksi, tetapi diharapkan mampu berkontribusi praktis dalam mendukung pengambilan keputusan berbasis data.

II. METODOLOGI PENELITIAN

Dalam penelitian ini, peneliti merancang metodologi yang digunakan untuk melakukan prediksi terhadap harga rumah. Pada penelitian ini peneliti menggunakan data sekunder yang berasal dari <https://www.kaggle.com/datasets/wisnuanggar/a/daftar-harga-rumah> yang diakses pada tanggal 16 Juli 2025 dengan populasi data seluruh harga rumah di Tebet yang terdiri dari 8 kolom dengan jumlah data sebanyak 1100 data. Pada penelitian ini sampel data diambil adalah harga rumah di wilayah Tebet.



Sumber: Penulis (2025)

Gambar 1. Metodologi Penelitian

Gambar 1 merupakan metodologi penelitian yang digunakan oleh peneliti sebagai pedoman dalam melakukan penelitian ini. Tahapan dalam penelitian ini dimulai dari pengumpulan data, dilanjutkan tahapan pre-processing yang termasuk didalamnya proses *data cleaning*, *feature selection* dan *splitting data*, kemudian tahapan dilanjutkan dengan pemodelan data menggunakan algoritma XGBRegressor (Al Khairi et al., 2025), selanjutnya model dievaluasi untuk mengukur tingkat akurasi menggunakan metrik regresi R2 Score, *Mean Absolute Error* (MAE) dan *Root Mean Squared Error* (RMSE). tahapan terakhir dilakukan analisis hasil pemodelan dengan data aktual untuk melihat pola kesalahan. Pada analisis ini peneliti melakukan analisis terhadap *feature important* yakni variabel yang paling berpengaruh terhadap harga rumah.

III. HASIL DAN PEMBAHASAN

Berdasarkan hasil penelitian yang dilakukan oleh peneliti terhadap dataset sekunder untuk dilakukan prediksi terhadap harga rumah, maka peneliti dapat menyajikan hasil proses pengolahan data, pelatihan model prediksi harga rumah menggunakan algoritma XGBRegressor, serta evaluasi terhadap kinerja model tersebut dalam pembahasan berikut:

1. Pengumpulan Data

Pada proses pengumpulan data ini peneliti menggunakan data sekunder yang bersumber dari kaggle dengan alamat website <https://www.kaggle.com/datasets/wisnuanggar/a/daftar-harga-rumah> yang diakses pada tanggal 16 Juli 2025 dengan populasi data seluruh harga rumah di Tebet yang terdiri dari 8

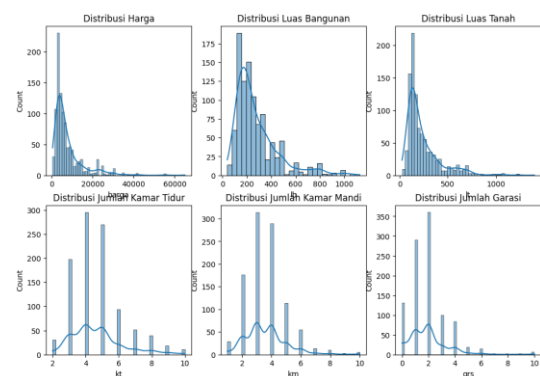
kolom dengan jumlah data sebanyak 1100 data. Data ini memiliki 8 fitur atau variabel diantaranya: nomor, nama rumah, harga rumah, luas bangunan, luas tanah, jumlah kamar tidur, jumlah kamar mandi, jumlah kapasitas mobil dalam garasi. Pada penelitian ini sampel data diambil adalah harga rumah di wilayah Tebet (Yuichi & Susetyo, 2025).

2. Pre-processing terdiri dari *data cleaning*, *feature selection*, *splitting data*

a. *Data cleaning*

Pada tahap *pre-processing* ini peneliti melakukan *data cleaning* yaitu menghilangkan variabel yang tidak diperlukan atau tidak berpengaruh terhadap variabel yang diteliti. Dimana pada tahapan ini peneliti menghapus nilai yang kosong dan menghilangkan variabel yang tidak diperlukan yaitu variabel nomor (Hakim, 2021)

Data yang sudah di bersihkan dapat dilihat pada grafik dimana berapa distribusi dari masing-masing variabel yang paling mendominasi dalam dataset.



Sumber: Penulis (2025)

Gambar 2. Grafik Distribusi Variabel dalam Dataset

Pada grafik diatas dapat dilihat bahwa dari variabel harga yang paling banyak mendominasi jumlah data rumah di Tebet adalah dari kisaran hingga ratusan juga 2 miliar, dari variabel luas bangunan yang paling banyak antara kisaran 100 sampai 300 meter, dari variabel luas tanah kisaran 100 meter, untuk variabel jumlah kamar tidur yang paling banyak adalah 4 kamar tidur, sedangkan variabel jumlah kamar mandi adalah 3, dan variabel kapasitas mobil dalam garasi yang paling banyak adalah kapasitas 2 mobil.

b. Feature Selection

Pada tahap *pre-processing* ini peneliti juga melakukan *feature selection* dengan memilih variabel yang relevan dengan harga rumah (Salwa & Malabay, 2025), dimana peneliti memilih variabel luas bangunan, luas tanah, jumlah kamar tidur, jumlah kamar mandi dan jumlah kapasitas mobil dalam garasi sebagai variabel yang paling memengaruhi harga rumah. Peneliti juga mengganti satuan harga agar lebih mudah dibaca dengan satuan harga per Rp 1000.000.

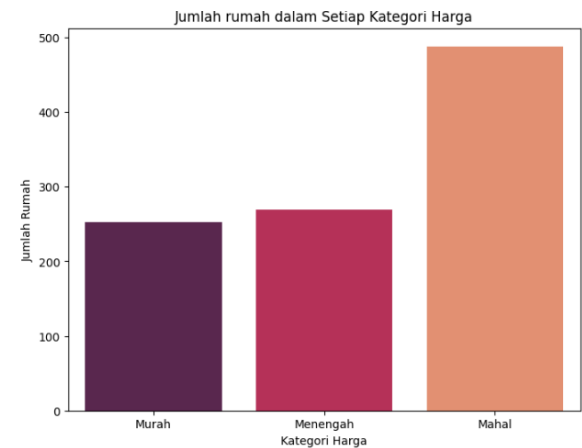
c. Splitting data

Pada tahap ini, peneliti membagi dataset ke dalam dua bagian (Nurmasani & Pristyanto, 2021), yaitu data latih (training data) dan data uji (testing data) dengan perbandingan 80:20. Sebanyak 80% dari keseluruhan data digunakan untuk melatih model agar dapat memahami pola hubungan antar variabel, sementara 20% sisanya digunakan untuk menguji kinerja model terhadap data yang belum pernah dilihat sebelumnya. Pembagian ini dilakukan secara acak (random split) untuk menghindari bias dan memastikan bahwa distribusi data pada masing-masing subset tetap representatif. Dengan cara ini, model dapat dievaluasi secara obyektif untuk mengetahui seberapa baik kemampuan prediksinya terhadap data baru. Langkah ini merupakan proses penting dalam machine learning karena memungkinkan peneliti untuk mengidentifikasi apakah model mengalami overfitting (terlalu baik pada data latih namun buruk pada data uji) atau underfitting (tidak cukup baik pada keduanya).

3. Klasifikasi

Pada tahapan ini peneliti melakukan klasifikasi harga rumah berdasarkan 3 kelas, yaitu Murah, Menengah dan Mahal. Peneliti membagi data harga rumah menjadi beberapa kelompok atau kelas berdasarkan quantile 1 (0.25) dan quantile 3 (0.75) dimana quantile 1 itu adalah harga rumah bernilai 340 juta, sedangkan quantile 3 itu harga rumah yang bernilai 6,5 miliar. Sehingga jika harga rumah berada di bawah $Q1 < 340$ juta dikategorikan sebagai Murah; dan jika harga rumah antara $Q1 < 340$ juta hingga $Q3 < 6,5$ miliar dikategorikan Menengah; dan di atas $Q3 > 6,5$ miliar dikategorikan sebagai Mahal. Dari hasil pembagian kelas ini menghasilkan 488 data yang masuk kategori kelas Murah, 269 data masuk kelas Menengah dan 253 data masuk kategori kelas Mahal. Tujuan peneliti membagi

data ini berdasarkan distribusi aktual bukan harga sembarangan dan berguna untuk klasifikasi harga rumah secara statistik dan data driven. Selain itu pembagian data ini juga bisa digunakan untuk memvisualisasikan kategori harga rumah sehingga mudah dibaca. Hasil klasifikasi harga rumah pada setiap kelasnya dapat dilihat pada gambar grafik berikut:



Sumber: Penulis (2025)

Gambar 3. Grafik Jumlah Rumah Dalam Setiap Kategori Harga

Selain dalam bentuk grafik peneliti juga menampilkan sample data hasil klasifikasi harga rumah dalam bentuk tabel seperti pada tabel berikut:

Tabel 1. Hasil Klasifikasi Harga Rumah

Nama_rumah	Harga	Lb	Lt	Kt	Km	Grs	Tingkat_harga
0 Rumah murah hook Tebet Timur, Jakarta Selatan	3800	220	220	3	3	0	Menengah
1 Rumah Modern di tebet dekat Stasiun, Tebet, Jakarta Selatan	4600	180	170	4	3	2	Menengah
2 Rumah Mewah 2 Lantai Hanya 3 Menit ke	3000	267	250	4	4	4	Murah

	Tebet, Tebet, Jakarta Selatan							
3	Rumah Baru	43 0	4 0	2 5	2	2	0	Murah
	Tebet, Tebet, Jakarta Selatan							
4	Rumah Bagus tebet komp Gudang Peluru It 350 m	90 00	4 0	3 5	6	5	3	Mahal

Sumber: Penulis (2025)

4. Pemodelan

Langkah selanjutnya peneliti melakukan proses pelatihan model menggunakan algoritma XGBRegressor. Algoritma ini dipilih oleh peneliti karena mampu menangani data non linear dalam jumlah besar dan kompleks (Juli et al., 2025), selain itu juga memiliki kinerja yang baik dalam memprediksi. Dalam tahap pemodelan ini peneliti melatih model pada data latih, kemudian menyesuaikan parameter model menggunakan teknik Ridge

5. Evaluasi model

Setelah model dibangun selanjutnya peneliti mengevaluasi model untuk mengukur akurasi menggunakan metrik regresi yang terdiri dari R2 Score (Coefficient of Determination), MAE (Mean Absolute Error). Dan MSE (Mean Squared Error) dan hasil evaluasi ini dapat digunakan untuk menilai seberapa baik model dalam memprediksi harga rumah.

Hasil evaluasi dari penelitian ini dapat terlihat pada tabel dibawah ini:

	Data Pelatihan	Data Pengujian
MAE	216.2706298828125	1668.155639648435
MSE	195273.3125	11507682.0
R2 Score	0.99648827314376	0.75349187850952

Sumber: Penulis (2025)

Berdasarkan tabel hasil evaluasi model, dapat dilihat bahwa nilai MAE pada data latih menghasilkan 216.2706298828125 yang berarti rata-rata selisih absolut antara nilai

prediksi dan nilai aktual lebih besar data pelatihan. Sedangkan nilai MAE pada data pengujian menghasilkan nilai yang lebih kecil yakni 1668.1556396484375.

Pada nilai MSE data pelatihan menghasilkan nilai 195273.3125 dan pada data pengujian 11507682.0 yang mana nilainya lebih kecil, sehingga hasil pengukuran kesalahan rata-rata kuadrat antara prediksi dan nilai aktual pada data pengujian lebih kecil dibandingkan dengan kesalahan rata-rata data pelatihan.

Hasil metrik evaluasi R2 score pada data pelatihan terlihat lebih baik dibandingkan dengan R2 score dari data pengujian. Hal ini berarti seberapa baik model menjelaskan variasi data target dibandingkan dengan model rata-rata. Dan pada data pelatihan hasilnya R2 score nya 0.9964882731437683 yang berarti model sangat baik menjelaskan 80-99% variasi data dan hasil dari data pengujian 0.7534918785095215 yang berarti model cukup baik menjelaskan sebagian besar variasi.

Berhubung hasil evaluasi model menghasilkan tingkat R2Score lebih baik pada data pelatihan dan hasilnya overfitting, maka peneliti melakukan evaluasi model Ridge untuk mengurangi overfitting, sehingga menghasilkan nilai R2 Score lebih baik pada data pengujian dengan nilai sebesar 0.7713084102181795 dan R2 score pada data pengujian sebesar 0.7040692834779283.

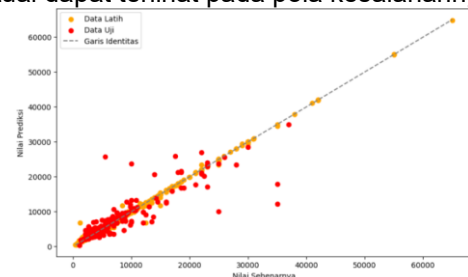
Untuk detail hasil evaluasi model Ridge dapat dilihat pada tabel dibawah:

	Data Pelatihan	Data Pengujian
MAE	2117.40699906882	1980.35625270868
MSE	16455645.4132114	10675957.0816483
R2 Score	0.70406928347792	0.77130841021817

Sumber: Penulis (2025)

6. Analisis hasil

Hasil prediksi dibandingkan dengan harga aktual dapat terlihat pada pola kesalahannya.

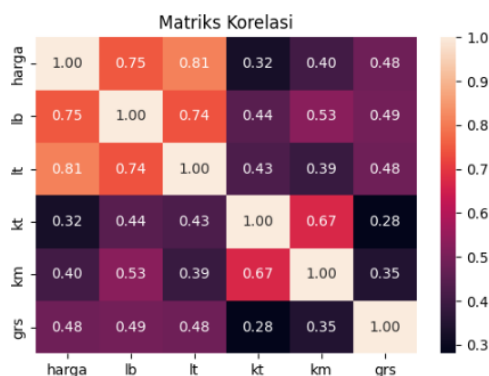


Sumber: Penulis (2025)

Gambar 4. Hasil Prediksi

Berdasarkan hasil prediksi dapat diketahui bahwa data latih dan data uji pada nilai 300000 tidak terlampau jauh selisih antara nilai prediksi dengan nilai sebenarnya atau harga aktual. Dan dapat dilihat bahwa hasil prediksi tidak meleset terlalu jauh dari harga aktualnya.

Dalam tahap analisis hasil ini, peneliti juga memvisualisasikan ke dalam matrik korelasi yang dilakukan terhadap variable yang memengaruhi harga rumah. Tujuannya adalah untuk mengetahui variable apa saja yang memiliki hubungan positif atau searah dengan variable yang akan di prediksi. Dalam tahap analisis korelasi ini peneliti mengukurnya dengan metode korelasi Pearson. Korelasi Pearson menghasilkan nilai antara -1 hingga 1, yang menunjukkan arah dan kekuatan hubungan antara dua variable (Fauzan Almahdy & Mega Pradnya, 2024).



Sumber: Penulis (2025)

Gambar 5. Matriks Korelasi

Dalam analisis korelasi ini peneliti menjadikan variable harga rumah sebagai acuan variable dependen. Berdasarkan grafik diatas dapat diketahui bahwa luas tanah memiliki nilai korelasi paling tinggi (0.81), kemudian disusul dengan variable luas bangunan dengan korelasi sebesar (0.75), kemudian variable garasi dengan nilai korelasi (0.48), diikuti variable kamar mandi (0.40) dan terakhir variable kamar tidur dengan nilai korelasi terendah (0.32).



Harga rumah impian anda diperkirakan sekitar IDR 7,339.389 juta

Sumber: Penulis (2025)

Gambar 6. Tampilan Interaktif Prediksi Harga Rumah Berdasarkan Fitur yang Dipilih

Pada gambar diatas pembaca dapat melihat prediksi harga rumah berdasarkan fitur yang dipilih, pembaca dapat memilih berdasarkan kebutuhan luas bangunan, luas tanah, jumlah kamar tidur, jumlah kamar mandi, dan kapasitas garasi sehingga pembaca dapat memperkirakan budget sesuai dengan kebutuhan rumah impian di masa depan. Sebagai contoh prediksi harga rumah sesuai hasil tampilan interaktif diatas, pembaca menginginkan rumah dengan luas bangunan 250 meter, luas tanah 300 meter, kamar tidur 3, kamar mandi 2 dan kapasitas garasi 2 maka perkiraan harga rumah 7.339.389 juta per meter.

IV. KESIMPULAN

Permasalahan kepemilikan rumah terutama bagi pasangan muda yang baru memulai kehidupan berkeluarga adalah harga rumah yang melonjak setiap tahunnya dan tidak sebanding dengan pendapatan masyarakat, sehingga penelitian ini dibuat untuk membangun model prediksi harga rumah menggunakan algoritma XGBRegressor yang merupakan metode machine learning berbasis ensemble dengan kemampuan generalisasi yang baik terhadap data kompleks. Model pada penelitian ini dilatih menggunakan data spesifikasi rumah yang mencakup berbagai variabel yang memengaruhi harga rumah seperti luas bangunan, luas tanah, jumlah kamar tidur, jumlah kamar mandi, dan kapasitas mobil dalam garasi. Hasil evaluasi pada penelitian ini menunjukkan bahwa model mampu memberikan tingkat akurasi yang cukup baik dengan nilai R2 Score yang tinggi dengan nilai 0.77 atau sekitar 77% serta nilai MAE dan MSE yang relatif rendah. Hal ini menandakan bahwa model yang digunakan dapat memprediksi harga rumah secara akurat. Secara praktis penelitian ini mampu memberikan kontribusi bagi pasangan muda yang ingin memiliki rumah pribadi untuk perencanaan finansial masa depan dalam bentuk alat bantu pengambilan keputusan berbasis data. Selain itu penelitian ini mampu menunjukkan bahwa penerapan machine learning dalam bidang properti juga mampu memberikan solusi untuk mendukung transparansi harga dan perencanaan finansial.

V. REFERENSI

Al Khairi, S., Adriansyah, A. R., & Rosyidi, L. (2025). Perbandingan XGB Regressor

- dengan Algoritma Lain untuk Prediksi Tarif Tol. *DBESTI: Journal of Digital Business and Technology Innovation*, 2(1), 127–132. <https://doi.org/10.54914/dbesti.v2i1.1477>
- Fauzan Almahdy, R., & Mega Pradnya, W. D. (2024). Prediksi Harga Rumah Di Kabupaten Bantul Menggunakan Algoritma Support Vector Regression. *Jurnal Teknik Informatika Dan Sistem Informasi*, 11(2), 152–165. <https://doi.org/10.35957/jatasi.v11i2.6001>
- Hakim, B. (2021). Analisa Sentimen Data Text Preprocessing Pada Data Mining Dengan Menggunakan Machine Learning. *JBASE - Journal of Business and Audit Information Systems*, 4(2), 16–22. <https://doi.org/10.30813/jbase.v4i2.3000>
- Hardjanto, K. R. (2025). (*Studi Kasus Di Kampung Intu Lingau , Kecamatan Nyuatan , Kabupaten Kutai Barat*) Kevin Renaldi Hardjanto *eJournal Pemerintahan Integratif Volume 11 , Nomor 4 , 2025*. 11(4), 436–446.
- Indonesia, P. R. (2024). Undang-Undang Republik Indonesia Nomor 3 Tahun 2024 Tentang Perubahan Kedua Atas Undang-Undang Nomor 6 Tahun 2014 Tentang Desa. *Peraturan .Bpk.Go.Id*, 31. <https://peraturan.bpk.go.id/Details/283617/uu-no-3-tahun-2024>
- Juli, V. N., Aulia, M. K., Utaminingsih, E., & Prihatin, N. (2025). *Computer Science (CO-SCIENCE) Model Prediksi Risiko Kesehatan Perkotaan Berbasis Lingkungan dengan XGBoost*. 5(2), 95–102.
- Mu'tashim, M. L., Muhayat, T., Damayanti, S. A., Zaki, H. N., & Wirawan, R. (2021). Analisis Prediksi Harga Rumah Sesuai Spesifikasi Menggunakan Multiple Linear Regression. *Informatik: Jurnal Ilmu Komputer*, 17(3), 238. <https://doi.org/10.52958/iftk.v17i3.3635>
- Nuris, N. (2024). Analisis Prediksi Harga Rumah Pada Machine Learning Metode Regresi Linear. *Explore*, 14(2), 108–112. <https://doi.org/10.35200/ex.v14i2.123>
- Nurmasani, A., & Pristyanto, Y. (2021). Algoritme Stacking Untuk Klasifikasi Penyakit Jantung Pada Dataset Imbalanced Class. *Pseudocode*, 8(1), 21–26. <https://doi.org/10.33369/pseudocode.8.1.21-26>
- Qolbi, M. B. S., Puteh, T. N., Rivandi, R., & Rozikin, C. (2025). Prediksi Harga Rumah Di Jakarta Pusat Menggunakan Algoritma Machine Learning. *Jurnal Ilmu Komputer Dan Bisnis*, 16(1), 16–24. <https://doi.org/10.47927/jikb.v16i1.840>
- Roeskamto, I. M., & Atmadja, A. S. (2022). Pergerakan Harga Rumah Di Indonesia: Analisis Sisi Permintaan. *Dimensi Utama Teknik Sipil*, 7(2), 53–59. <https://doi.org/10.9744/duts.7.2.53-59>
- Rustini, T., Indah Sari, D. N., Zahidah Fadhilah, D. N., & Husna Fadhilah, S. D. (2025). Kondisi Ketika Kebutuhan Sekunder dan Tersier Lebih Diutamakan dibanding Kebutuhan Primer. *Jurnal Pendidikan Tambusai*, 9(1), 3045–3052.
- Salwa, M. A., & Malabay. (2025). Optimasi Model Algoritma Klasifikasi Data Mining Menggunakan Metode Feature Selection Untuk Prediksi Penyakit Stroke. *Tekinfor*, 26(1), 11–20.
- Score, A. I. (2024). *Mengapa Harga Rumah Terus Naik? Ini Faktor dan Dampaknya*. <https://www.idscore.id/articles/mengapa-harga-rumah-terus-naik-ini-faktor-dan-dampaknya>
- Warjiyono, Rais, A. N., Alfaro, I., Hadi, S. W., & Kurniawan, W. (2024). Analisa Prediksi Harga Jual Rumah Menggunakan Algoritma Random Forest Machine Learning. *JURSI STEKNI (Jurnal Sistem Informasi Dan Teknologi Informasi)*, 6(2), 416–423.
- Wicaksana, A. P., Christioko, V. B., Musyaffa, F. I., Saputro, R. E., & Mahendra, R. (2024). SEMINAR NASIONAL INOVASI DAN TREN TEKNOLOGI (SINATTI) Fakultas Teknologi Informasi dan Komunikasi Universitas Semarang Analisis Regresi Linear untuk Prediksi Harga Rumah Berdasarkan Luas Area Tanah Menggunakan Dataset Kaggle. *Universitas Semarang*, 2, 1–10.
- Yuichi, M., & Susetyo, Y. A. (2025). Klasifikasi Penyakit Migrain dengan Metode Naïve Bayes pada Dataset Kaggle. *Jurnal Indonesia: Manajemen Informatika Dan Komunikasi*, 6(1), 139–151. <https://doi.org/10.35870/jimik.v6i1.1150>