

Analisis Prediksi Harga Rumah di Ames, Iowa

Menggunakan Pendekatan Stacking Regressor

Abdurrahman Assegaf | 11 Juli 2025



Latar Belakang

Proyek ini bertujuan memprediksi harga rumah menggunakan dataset Ames Housing, mencakup **80 fitur** seperti luas kavling, kualitas bangunan, dan jumlah kamar.

Prediksi harga rumah sangat krusial dalam pasar properti untuk **investor**, **pembeli**, dan **penjual** guna mengambil keputusan yang tepat.

Akurasi model dapat memberikan wawasan pasar yang **berharga** dan **mengurangi risiko** finansial.



Tujuan Proyek



Membangun Model Akurat

Membangun model machine learning untuk memprediksi harga rumah dengan akurasi tinggi menggunakan pendekatan **stacking regressor**.



Menghasilkan Prediksi Data Uji

Menghasilkan prediksi untuk data uji dan menyimpannya dalam file **submission.csv** yang siap untuk kompetisi.

Metodologi

Langkah-langkah Utama



Pemuatan & Pra-pemrosesan Data

- Penanganan nilai hilang
- Pengkodean fitur kategorikal
- Normalisasi data

Pengembangan Model Stacking Regressor

- GradientBoosting
- XGBoost
- LGBM
- Ridge



Evaluasi Model

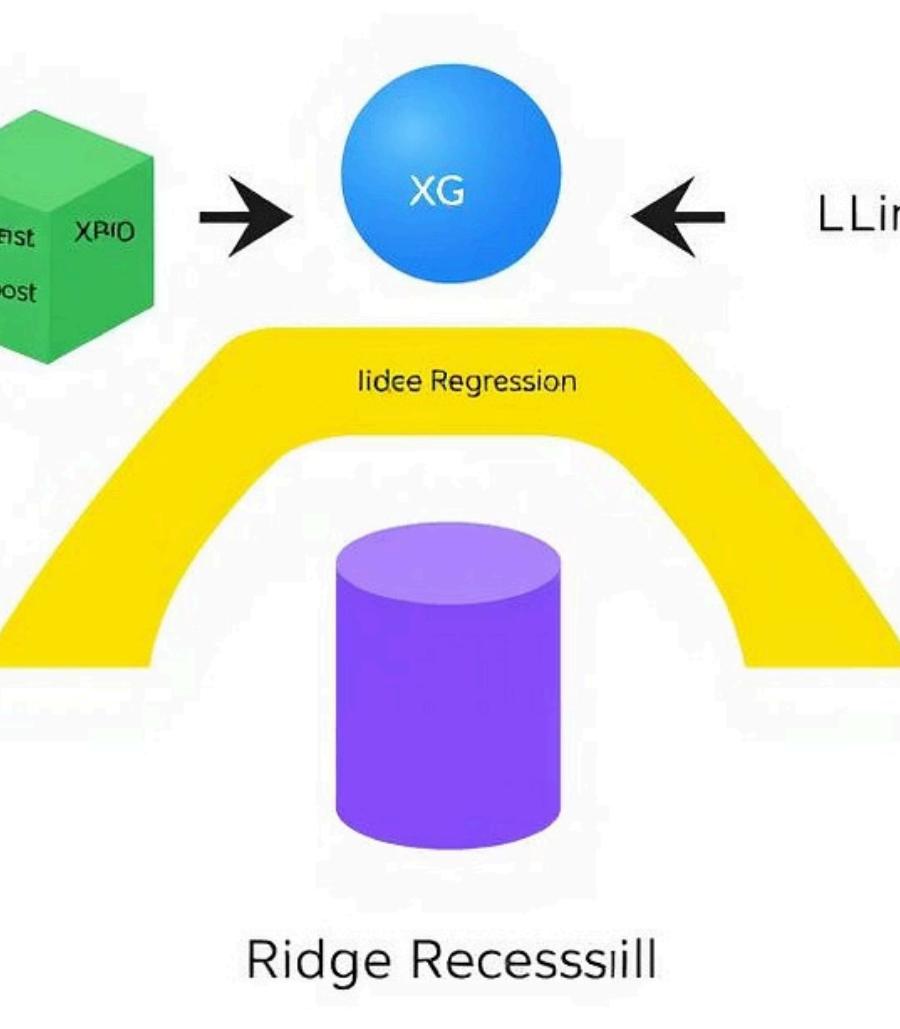
- Menggunakan metrik RMSE

Prediksi & Finalisasi

- Prediksi pada data uji
- Pembuatan submission.csv

Model Stacking Regressor

Model ini mengombinasikan kekuatan beberapa model dasar untuk meningkatkan akurasi prediksi.



Model Dasar (Base Models)

- **GradientBoostingRegressor:**
(learning_rate=0.01,
max_depth=15,
n_estimators=1000)
- **XGBRegressor:**
(learning_rate=0.1,
max_depth=3,
n_estimators=300)
- **LGBMRegressor:** (parameter
disesuaikan untuk performa
optimal)

Estimator Akhir (Meta-Estimator)

- **Ridge Regressor:** Digunakan sebagai model akhir dengan bobot [2, 3, 1] untuk mengombinasikan output model dasar.
- Pendekatan ini memanfaatkan kekuatan masing-masing model.

Hasil dan Evaluasi

Metrik Kinerja

Nilai **RMSE** (Root Mean Squared Error) yang diperoleh adalah **0.12093** pada skala logaritmik.

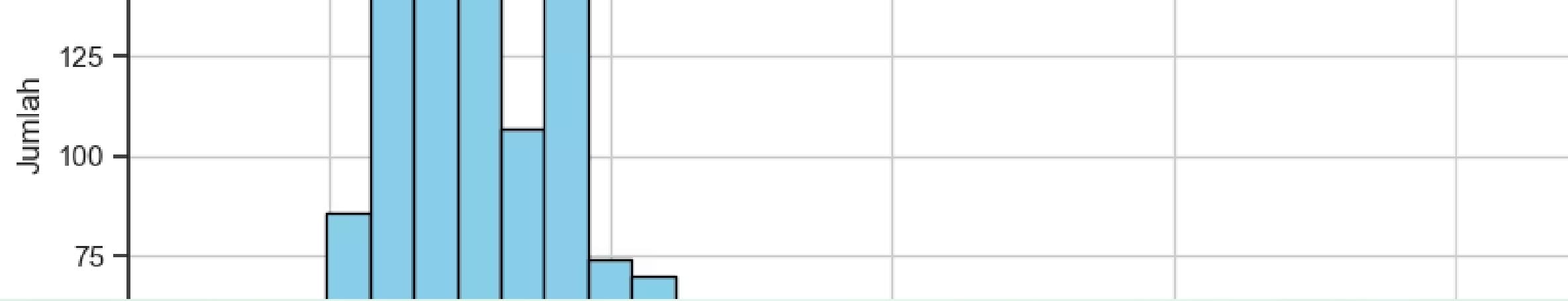
Nilai RMSE yang rendah ini menunjukkan bahwa model kami memiliki akurasi yang **sangat baik** dalam memprediksi harga rumah.

Prediksi akhir untuk 1459 entri data uji telah disimpan dalam file **submission.csv**.

Contoh Prediksi Harga

Id Properti	Harga Prediksi (IDR)
1461	Rp 1.850.000.000
2919	Rp 3.400.000.000

Rentang harga prediksi dari model kami berkisar dari **Rp 770 Juta** hingga **Rp 7,7 Miliar**, menunjukkan kemampuan model dalam menangani variasi harga yang luas.



Visualisasi Hasil

Distribusi harga prediksi menunjukkan pola yang mendekati distribusi normal, menandakan model mampu menangkap variasi pasar secara efektif.

Fitur Paling Berpengaruh

- Berdasarkan analisis, fitur "**Kualitas Keseluruhan**" dan "**Luas Area Tinggal**" adalah prediktor paling signifikan terhadap harga rumah.

Kesimpulan

- **Efektivitas Model:** Model stacking regressor berhasil memprediksi harga rumah dengan **RMSE 0.12093**, menunjukkan akurasi tinggi.
- **Penanganan Dataset Kompleks:** Pendekatan ini terbukti efektif dalam menangani dataset yang kompleks dengan banyak fitur.
- **Siap untuk Kompetisi:** Prediksi telah disimpan dalam **submission.csv**, siap untuk digunakan dalam kompetisi atau aplikasi nyata.

Saran untuk Pengembangan



Rekayasa Fitur Lanjutan

Menciptakan fitur baru seperti usia rumah (Tahun Dibangun - Tahun Dijual), total kamar mandi, atau total area lantai.



Penyesuaian Hiperparameter Optimal

Menggunakan teknik seperti **GridSearchCV** atau **RandomizedSearchCV** untuk menemukan kombinasi hiperparameter terbaik.



Validasi Silang yang Robust

Menerapkan strategi validasi silang yang lebih robust untuk estimasi performa model yang lebih andal dan generalisasi yang baik.

Terima Kasih atas perhatian Anda!

Ada pertanyaan?