

Implementasi Regresi Linear Berganda dengan Metode Normal Equation untuk Prediksi Harga Properti

Muhamad Dimas Saputra

Teknik Informatika/Sains dan Teknologi

UNIVERSITAS ISLAM NEGERI SYARIF HIDAYATULLAH JAKARTA

Tangerang Selatan, Indonesia

muhamad.dimas24@mhs.uinjkt.ac.id

Abstract—Sektor properti memiliki peran vital dalam ekonomi, namun penentuan harga rumah seringkali menjadi tantangan kompleks karena dipengaruhi oleh banyak faktor fisik. Penelitian ini bertujuan untuk mengimplementasikan metode Regresi Linear Berganda menggunakan pendekatan *Normal Equation* untuk memprediksi harga rumah di Jakarta Selatan. Berbeda dengan pendekatan iteratif seperti *Gradient Descent*, *Normal Equation* memberikan solusi analitik langsung untuk menemukan parameter model yang optimal tanpa memerlukan penentuan *learning rate* atau iterasi berulang. Data yang digunakan mencakup fitur luas tanah, luas bangunan, jumlah kamar tidur, jumlah kamar mandi, dan ketersediaan garasi. Hasil pengujian menunjukkan bahwa model mampu menghasilkan prediksi dengan nilai *R-squared* (R^2) sebesar 0.836 (83.6%), yang mengindikasikan bahwa model memiliki performa sangat baik dalam menjelaskan variasi harga properti. Tingkat kesalahan prediksi yang diukur dengan *Root Mean Squared Error* (RMSE) adalah sebesar 3.97 miliar rupiah.

Index Terms—Prediksi Harga Rumah, Regresi Linear, Normal Equation, Jakarta Selatan, Machine Learning.

I. PENDAHULUAN

A. Latar Belakang

Sektor properti memegang peranan vital dalam perekonomian global. Rumah tidak hanya berfungsi sebagai kebutuhan pokok, tetapi juga instrumen investasi yang strategis karena nilainya yang cenderung meningkat [?]. Namun, memprediksi harga properti secara akurat merupakan tantangan tersendiri karena dipengaruhi oleh berbagai variabel fisik seperti luas bangunan, jumlah kamar, dan fasilitas pendukung lainnya [?].

Pendekatan *Machine Learning* menawarkan solusi untuk masalah prediksi ini. Salah satu metode yang efektif adalah Regresi Linear Berganda, yang memodelkan hubungan linear antara variabel independen (fitur rumah) dan variabel dependen (harga). Penelitian terdahulu menunjukkan bahwa metode ini mampu memberikan estimasi harga yang dapat diandalkan [?], [?].

Dalam penelitian ini, optimasi model dilakukan menggunakan metode *Normal Equation*. Berbeda dengan algoritma *Gradient Descent* yang memperbarui parameter secara iteratif dan memerlukan penyetelan *hyperparameter* seperti *learning rate*, *Normal Equation* menyelesaikan masalah *Least Squares* secara analitik menggunakan operasi matriks. Metode ini

efisien untuk dataset dengan jumlah fitur yang tidak terlalu besar, seperti dataset harga rumah Jakarta Selatan yang digunakan dalam studi ini.

B. Rumusan Masalah

Berdasarkan latar belakang tersebut, rumusan masalah penelitian ini adalah:

- 1) Bagaimana penerapan metode *Normal Equation* dalam Regresi Linear Berganda untuk memprediksi harga rumah?
- 2) Seberapa besar pengaruh fitur fisik (Luas Tanah, Luas Bangunan, Jumlah Kamar, Garasi) terhadap harga rumah di Jakarta Selatan?
- 3) Bagaimana kinerja model yang dihasilkan jika diukur dengan metrik evaluasi R^2 dan RMSE?

C. Tujuan Penelitian

Tujuan dari penelitian ini adalah:

- 1) Mengimplementasikan algoritma Regresi Linear Berganda dengan solusi analitik *Normal Equation*.
- 2) Membangun model prediksi harga rumah berdasarkan dataset properti Jakarta Selatan.
- 3) Mengevaluasi akurasi model untuk mengetahui tingkat kesalahan prediksi dan kecocokan model terhadap data.

D. Batasan Masalah

- **Metode:** Menggunakan *Normal Equation* untuk mencari parameter bobot optimal θ .
- **Dataset:** Data harga rumah di Jakarta Selatan dengan fitur: Luas Tanah (LT), Luas Bangunan (LB), Jumlah Kamar Tidur (JKT), Jumlah Kamar Mandi (JKM), dan Garasi (GRS).
- **Evaluasi:** Kinerja model diukur menggunakan MSE, RMSE, dan R^2 Score.

II. LANDASAN TEORI

A. Regresi Linear Berganda

Regresi Linear Berganda memodelkan hubungan antara satu variabel dependen Y (harga) dengan beberapa variabel inde-

penden x_1, x_2, \dots, x_n (fitur). Persamaan model didefinisikan sebagai:

$$Y = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \dots + \theta_n x_n \quad (1)$$

Dimana θ adalah parameter bobot yang meminimalkan fungsi biaya.

B. Normal Equation

Normal Equation adalah metode untuk menemukan nilai θ yang meminimalkan *Cost Function* $J(\theta)$ secara langsung dengan menetapkan turunannya menjadi nol. Solusi analitiknya diberikan oleh persamaan:

$$\theta = (X^T X)^{-1} X^T y \quad (2)$$

Dimana X adalah matriks desain (termasuk kolom bias) dan y adalah vektor target. Metode ini memberikan solusi eksak untuk masalah *linear least squares* [?].

III. METODOLOGI

A. Pra-pemrosesan Data

Data mentah melalui beberapa tahapan pra-pemrosesan sebelum pelatihan:

- 1) **Pembersihan Data:** Kolom yang tidak relevan seperti 'KOTA' (karena nilainya seragam 'JAKSEL') dihapus.
- 2) **Encoding:** Fitur kategorikal 'GRS' (Garasi) diubah menjadi format numerik (1 untuk 'ADA', 0 untuk 'TIDAK ADA').
- 3) **Pembagian Data:** Dataset dibagi menjadi data latih (90%) dan data uji (10%) untuk validasi model.

B. Implementasi

Model diimplementasikan menggunakan Python dengan pustaka NumPy. Fungsi *Normal Equation* dibangun dari awal (*from scratch*) untuk menghitung parameter θ menggunakan operasi perkalian matriks dan inversi, tanpa menggunakan fungsi *black-box* dari pustaka *machine learning* tingkat tinggi.

IV. HASIL DAN PEMBAHASAN

A. Evaluasi Model

Model dilatih menggunakan data latih dan diuji pada data uji yang belum pernah dilihat sebelumnya. Hasil evaluasi kinerja model adalah sebagai berikut:

- **Mean Squared Error (MSE):** 1.57×10^{19}
- **Root Mean Squared Error (RMSE):** Rp 3.971.501.609
- **R-squared (R^2):** 0.8365 (83.65%)
- **Mean Absolute Percentage Error (MAPE):** 27.08%

B. Analisis

Nilai R^2 sebesar 0.8365 menunjukkan bahwa model mampu menjelaskan sekitar 83.6% variasi harga rumah di Jakarta Selatan berdasarkan fitur-fitur yang digunakan. Ini menunjukkan korelasi yang kuat antara luas tanah/bangunan dengan harga jual. Nilai RMSE sebesar 3.97 miliar rupiah, meskipun terlihat besar secara nominal, cukup wajar mengingat varians harga properti di kawasan elite Jakarta Selatan yang sangat tinggi (mencapai puluhan miliar rupiah). Penggunaan *Normal Equation* terbukti efektif dan cepat untuk jumlah data ini.

V. KESIMPULAN

Penelitian ini berhasil menerapkan metode *Normal Equation* untuk prediksi harga rumah. Model menghasilkan tingkat akurasi yang baik dengan R^2 di atas 80%. Pendekatan analitik ini memberikan solusi parameter optimal secara efisien tanpa memerlukan proses iterasi pelatihan.

REFERENCES