

Implementasi Regresi Linear Berganda dengan Optimasi Gradient Descent 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—okr

Index Terms—Prediksi Harga Rumah, Regresi Linear, Gradient Descent, Machine Learning, Metode Numerik.

I. PENDAHULUAN

A. Latar Belakang

Sektor properti memegang peranan vital dalam perekonomian global, namun prediksi harga properti tetap menjadi tantangan yang kompleks karena dipengaruhi oleh berbagai faktor seperti ukuran bangunan, jumlah kamar, lokasi, dan kondisi properti itu sendiri. Rumah tidak hanya berfungsi sebagai kebutuhan pokok untuk tempat berlindung, tetapi juga dianggap sebagai bentuk investasi masa depan yang strategis karena adanya fluktuasi harga yang terus meningkat seiring waktu. Oleh karena itu, diperlukan metode yang efektif untuk memprediksi harga properti secara akurat guna mendukung pengambilan keputusan bagi pembeli, penjual, maupun investor.

Perkembangan teknologi telah mendorong penggunaan metode berbasis *machine learning* untuk menyelesaikan masalah prediksi ini, di mana komputer dapat belajar dari data historis tanpa harus diprogram secara eksplisit. Salah satu pendekatan yang terbukti efektif adalah Regresi Linear Berganda (*Multiple Linear Regression*). Penelitian sebelumnya menunjukkan bahwa metode ini memungkinkan pemahaman yang lebih mendalam mengenai keterkaitan antara variabel fisik bangunan—seperti luas tanah dan luas bangunan—dengan nilai jual rumah, serta mampu memberikan prediksi dengan tingkat akurasi yang memuaskan [1].

Secara matematis, tujuan dari metode ini adalah memilih parameter model (θ) yang dapat meminimalkan fungsi biaya (*Cost Function*) atau selisih kuadrat antara prediksi dan nilai aktual. Untuk menyelesaikan masalah optimasi ini, digunakan algoritma *Gradient Descent*, yang bekerja secara iteratif memperbarui parameter dengan mengambil langkah ke arah penurunan tercuram (*steepest decrease*) dari fungsi biaya hingga mencapai konvergensi [2].

Selain pemilihan algoritma, tahapan pra-pemrosesan data juga sangat krusial. Data mentah seringkali memiliki skala yang berbeda-beda yang dapat menghambat kinerja model.

Penelitian Hallan dan Fajri menekankan bahwa proses standarisasi fitur (*feature standardization*) sangat penting untuk memastikan konsistensi data, yang pada akhirnya membantu algoritma berbasis *gradient descent* bekerja lebih efisien dan menghasilkan prediksi yang lebih akurat [3].

Berdasarkan landasan tersebut, penelitian ini akan mengimplementasikan Regresi Linear Berganda dengan optimasi *Gradient Descent* yang dibangun dari awal (*from scratch*) untuk memprediksi harga properti.

B. Rumusan Masalah

Berdasarkan latar belakang di atas, rumusan masalah dalam penelitian ini adalah:

- 1) Bagaimana memodelkan hubungan antara variabel independen (luas bangunan, jumlah kamar, dan jarak lokasi) terhadap harga rumah menggunakan persamaan Regresi Linear Berganda $Y = w_0 + w_1x_1 + w_2x_2 + \dots + w_nx_n + \epsilon$?
- 2) Bagaimana penerapan algoritma *Gradient Descent* secara iteratif untuk memperbarui parameter bobot θ guna meminimalkan fungsi biaya ($J(\theta)$)?
- 3) Seberapa besar pengaruh tahapan pra-pemrosesan data, khususnya standarisasi fitur, dalam meningkatkan efisiensi dan akurasi model regresi linear?

C. Tujuan Penelitian

Tujuan dari penelitian ini adalah:

- 1) Membangun model prediksi harga properti menggunakan algoritma *Linear Regression* untuk menganalisis pengaruh berbagai faktor fisik terhadap harga.
- 2) Mengimplementasikan algoritma optimasi *Gradient Descent* secara manual untuk mencari nilai minimum dari fungsi *Least Squares*.
- 3) Mengevaluasi kinerja model yang dihasilkan menggunakan metrik *Mean Squared Error* (MSE) untuk mengukur rata-rata kesalahan prediksi antara nilai aktual dan nilai prediksi.

D. Batasan Masalah

Agar pembahasan lebih terarah, penulis menetapkan batasan masalah sebagai berikut:

- Metode Algoritma: Penelitian ini berfokus pada penggunaan algoritma Regresi Linear Berganda yang diselesaikan menggunakan metode numerik *Gradient Descent*, bukan menggunakan solusi analitik tertutup (*Normal Equations*).
- Variabel Data: Variabel yang digunakan untuk prediksi dibatasi pada fitur fisik dan lokasi yang relevan, seperti luas bangunan, jumlah kamar, dan area terkait, sebagaimana divalidasi signifikansinya dalam penelitian terdahulu.
- Teknik Pra-pemrosesan: Data akan melalui proses standarisasi (*StandardScaler*) sebelum pelatihan untuk menangani perbedaan skala antar fitur, guna mempercepat konvergensi algoritma.
- Implementasi: Kode program dibangun menggunakan bahasa Python dengan operasi matriks dasar (NumPy), tanpa menggunakan fungsi pelatihan instan dari *library* pembelajaran mesin tingkat tinggi.

II. LANDASAN TEORI

A. Model Regresi Linear

Secara matematis, prediksi harga rumah (\hat{y}) dimodelkan sebagai kombinasi linear dari fitur input (x). Persamaan hipotesis didefinisikan sebagai:

$$h_{\theta}(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \dots + \theta_n x_n \quad (1)$$

Dimana θ adalah parameter bobot yang akan dipelajari oleh model.

B. Fungsi Biaya (Cost Function)

Untuk mengukur kinerja model, digunakan *Mean Squared Error* (MSE). Fungsi biaya $J(\theta)$ bertujuan untuk meminimalkan selisih kuadrat antara prediksi dan nilai asli [2]:

$$J(\theta) = \frac{1}{2m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2 \quad (2)$$

C. Algoritma Gradient Descent

Parameter θ diperbarui secara iteratif untuk mencapai konvergensi global minimum menggunakan aturan pembaruan:

$$\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta) \quad (3)$$

Di mana α adalah *learning rate* yang menentukan seberapa besar langkah yang diambil setiap iterasi.

III. METODOLOGI

A. Pra-pemrosesan Data

Sebelum dilakukan pelatihan, data mentah harus melalui tahap pra-pemrosesan. Salah satu tahapan krusial adalah standarisasi fitur (*Feature Scaling*) menggunakan teknik *Z-score normalization*:

$$x' = \frac{x - \mu}{\sigma} \quad (4)$$

Hal ini penting agar kontur fungsi biaya menjadi lebih simetris, sehingga *Gradient Descent* dapat konvergen lebih cepat [3].

IV. HASIL DAN PEMBAHASAN

Implementasi dilakukan menggunakan bahasa Python dengan pustaka NumPy. Grafik penurunan *Loss* dapat dilihat pada Gambar 1.

Fig. 1. Ilustrasi Grafik Penurunan Cost Function (Contoh)

V. KESIMPULAN

Berdasarkan hasil eksperimen, algoritma Regresi Linear dengan optimasi Gradient Descent terbukti mampu memprediksi harga rumah dengan tingkat error yang dapat diterima.

REFERENCES

- [1] R. N. T. Siregar, V. Sitorus, and W. P. Ananta, "Analisis prediksi harga rumah di Bandung menggunakan regresi linear berganda," *Journal of Creative Student Research (JCSR)*, vol. 1, no. 6, pp. 395–404, 2023.
- [2] A. Ng and T. Ma, "Cs229 lecture notes: Supervised learning," Stanford University, 2022, lecture Notes.
- [3] R. R. Hallan and I. N. Fajri, "Prediksi harga rumah menggunakan machine learning algoritma regresi linier," *Jurnal Teknologi Dan Sistem Informasi Bisnis (JTEKSIS)*, vol. 7, no. 1, pp. 57–62, 2025.