

SUPERVISED MACHINE LEARNING: REGRESSION

Nama : Ailsa Najwa Insyira Putri

NIM : A11.2023.14860

1. Pendahuluan

Supervised Machine Learning merupakan salah satu pendekatan dalam machine learning yang digunakan untuk membangun model prediktif berdasarkan data berlabel. Salah satu permasalahan yang umum diselesaikan dengan pendekatan ini adalah **prediksi nilai numerik**, seperti harga rumah.

Pada tugas akhir ini, dilakukan pembangunan **model regresi** untuk memprediksi harga rumah berdasarkan beberapa fitur numerik, yaitu luas bangunan (Area), jumlah kamar tidur (Bedrooms), dan jumlah kamar mandi (Bathrooms). Tujuan dari tugas ini adalah untuk memahami hubungan antara fitur-fitur tersebut dengan harga rumah serta mengevaluasi performa model regresi yang dibangun sebelum digunakan pada permasalahan yang lebih kompleks.

2. Dataset dan Metodologi

Dataset yang digunakan dalam penelitian ini terdiri dari **300 data rumah** dengan empat variabel utama, yaitu Area, Bedrooms, Bathrooms, dan Price sebagai variabel target. Dataset ini merepresentasikan karakteristik fisik rumah yang umum digunakan dalam penentuan harga properti.

Tahapan metodologi yang dilakukan meliputi:

1. **Eksplorasi data awal** untuk memahami distribusi dan karakteristik data.
2. **Pemisahan data** menjadi data latih dan data uji dengan rasio 80:20.
3. **Pembangunan model** menggunakan algoritma **Linear Regression**.
4. **Evaluasi performa model** menggunakan metrik Mean Absolute Error (MAE), Mean Squared Error (MSE), dan R^2 Score.

Pendekatan ini dipilih karena Linear Regression merupakan metode dasar yang efektif untuk memodelkan hubungan linier antara variabel independen dan variabel target.

3. Hasil dan Evaluasi Model

Hasil pelatihan model menunjukkan bahwa algoritma Linear Regression mampu mempelajari hubungan antara fitur-fitur numerik dengan harga rumah. Evaluasi performa model dilakukan menggunakan tiga metrik utama, yaitu MAE, MSE, dan R^2 Score.

Nilai MAE dan MSE menunjukkan tingkat kesalahan prediksi yang masih dapat diterima, sementara nilai R^2 Score menunjukkan bahwa sebagian besar variasi harga rumah dapat dijelaskan oleh fitur Area, Bedrooms, dan Bathrooms. Hal ini mengindikasikan bahwa model memiliki kemampuan prediksi yang cukup baik dan dapat digunakan sebagai dasar untuk pengembangan model yang lebih kompleks.

4. Kesimpulan

Berdasarkan hasil penelitian, dapat disimpulkan bahwa model regresi yang dibangun berhasil memprediksi harga rumah berdasarkan fitur-fitur numerik yang digunakan. Luas bangunan, jumlah kamar tidur, dan jumlah kamar mandi terbukti memiliki pengaruh terhadap harga rumah.

Meskipun model Linear Regression memberikan hasil yang cukup baik, pengembangan lebih lanjut masih dapat dilakukan, seperti menambahkan fitur lain (misalnya lokasi atau usia bangunan) serta menggunakan algoritma regresi yang lebih kompleks untuk meningkatkan akurasi prediksi.