

**Laporan tugas akhir mata kuliah kecerdasan buatan**

Judul : Prediksi harga rumah di jaksel

Kelompok 3



Disusun oleh :

Mhd nurdin al kahfi

Sabri mutiur rahman

Rafii abdurrahman

**Fakultas sains dan teknologi, program studi teknik informatika**

**Universitas darussalam gontor**

# 1. Pendahuluan

## 1.1 Latar Belakang

Dalam beberapa tahun terakhir, industri real estat di Tebet, Indonesia, telah mengalami pertumbuhan yang pesat, menjadikannya area yang menarik untuk penelitian dan eksperimen. Salah satu tugas umum dalam bidang ini adalah meramalkan harga rumah, yang bisa sulit karena sifat pasar yang kompleks dan dinamis. Dalam eksperimen ini, kami akan menggunakan regresi linear, sebuah metode statistik yang banyak digunakan, untuk menganalisis data dan membuat prediksi tentang harga rumah di Tebet. Dengan menerapkan teknik ini, kami bertujuan untuk mendapatkan wawasan tentang faktor-faktor yang memengaruhi harga rumah dan meningkatkan kemampuan kami untuk meramalkannya secara akurat.

## 1.2 Tujuan

Penelitian ini bertujuan untuk membangun model prediksi harga rumah menggunakan regresi linear berdasarkan data rumah di Tebet, Jakarta Selatan.

## 1.3 Metode yang Digunakan

Penelitian ini menggunakan metode regresi linear untuk menganalisis hubungan antara variabel independen (luas bangunan, luas tanah, jumlah kamar tidur, jumlah kamar mandi, dan kapasitas garasi) terhadap variabel dependen (harga rumah).

# 2. Tinjauan Pustaka

## 2.1 Konsep Regresi Linear

Regresi linear merupakan salah satu metode statistik yang digunakan untuk menganalisis hubungan antara satu atau lebih variabel independen dengan variabel dependen.

Tujuan dari regresi linear adalah untuk menemukan koefisien  $\beta$  yang meminimalkan kesalahan kuadrat rata-rata antara nilai yang diamati dan nilai yang diprediksi oleh model.

Pros:

Simpel dan Mudah diinterpretasi: Regresi linear adalah salah satu metode yang paling sederhana dalam statistik yang mudah dipahami dan diinterpretasikan.

Fleksibel: Regresi linear dapat dengan mudah diperluas untuk mempertimbangkan hubungan non-linear antara variabel dependen dan independen dengan menggunakan transformasi variabel.

Penggunaan yang Luas: Regresi linear banyak digunakan dalam berbagai bidang seperti ekonomi, ilmu sosial, kedokteran, dan lain-lain karena kemampuannya untuk memberikan insight tentang hubungan antara variabel.

Cons:

Asumsi yang Kuat: Regresi linear memiliki beberapa asumsi yang harus dipenuhi, termasuk hubungan linear antara variabel dependen dan independen, homoskedastisitas, dan independensi dari kesalahan.

Sensitif terhadap Outliers: Regresi linear rentan terhadap pengaruh outliers, yang dapat mengganggu estimasi parameter dan keakuratan prediksi.

Keterbatasan dalam Menangani Variabel Kategorikal: Regresi linear memiliki keterbatasan dalam menangani variabel kategorikal tanpa transformasi tambahan seperti one-hot encoding, yang dapat mempengaruhi interpretasi koefisien.

Evaluasi Model pada Data Pelatihan:

MAE: 2117.4473452271177

MSE: 16455644.894573443

R2 Score: 0.7040692928048742

Evaluasi Model pada Data Pengujian:

MAE: 1980.3555183722162

MSE: 10675731.022166723

R2 Score: 0.7713132526788451

## 2.2 Faktor-faktor yang Mempengaruhi Harga Rumah

Beberapa faktor utama yang mempengaruhi harga rumah meliputi luas tanah, luas bangunan, jumlah kamar tidur, jumlah kamar mandi, serta fasilitas tambahan seperti garasi.

## 3. Metodologi

### 3.1 Data yang Digunakan

Dataset diambil dari kaggle

Dataset yang digunakan berasal dari data rumah di Tebet, Jakarta Selatan dengan atribut sebagai berikut:

nomor : nomor data

nama\_rumah : nama rumah

harga : harga rumah

lb : luas bangunan

lt : luas tanah

kt : jumlah kamar tidur

km : jumlah kamar mandi

grs : kapasitas mobil dalam garasi

### 3.2 Proses Pengolahan Data

Pembersihan Data: Menghapus data yang duplikat dan menangani nilai yang hilang.

Pembagian Data: Dataset dibagi menjadi data latih dan data uji.

Implementasi Model: Model regresi linear diterapkan menggunakan pustaka sklearn.

## 4. Hasil dan Pembahasan

### 4.1 Analisis Deskriptif

Analisis statistik dilakukan untuk memahami distribusi data, nilai rata-rata, dan outlier dalam dataset.

### 4.2 Evaluasi Model

Model dievaluasi menggunakan metrik berikut:

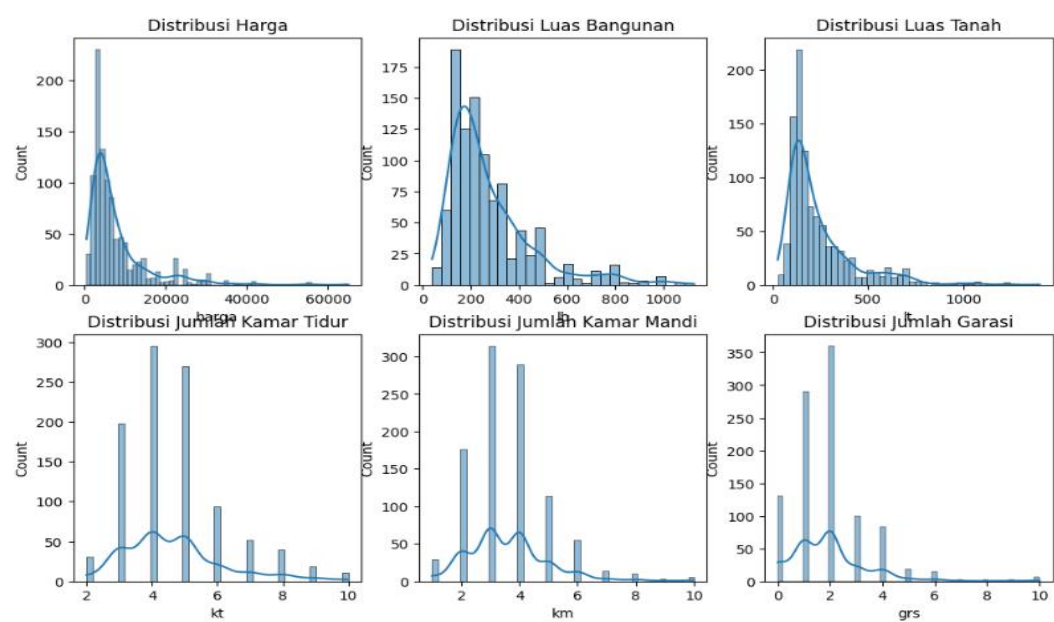
Mean Squared Error (MSE)

R-squared ( $R^2$ )

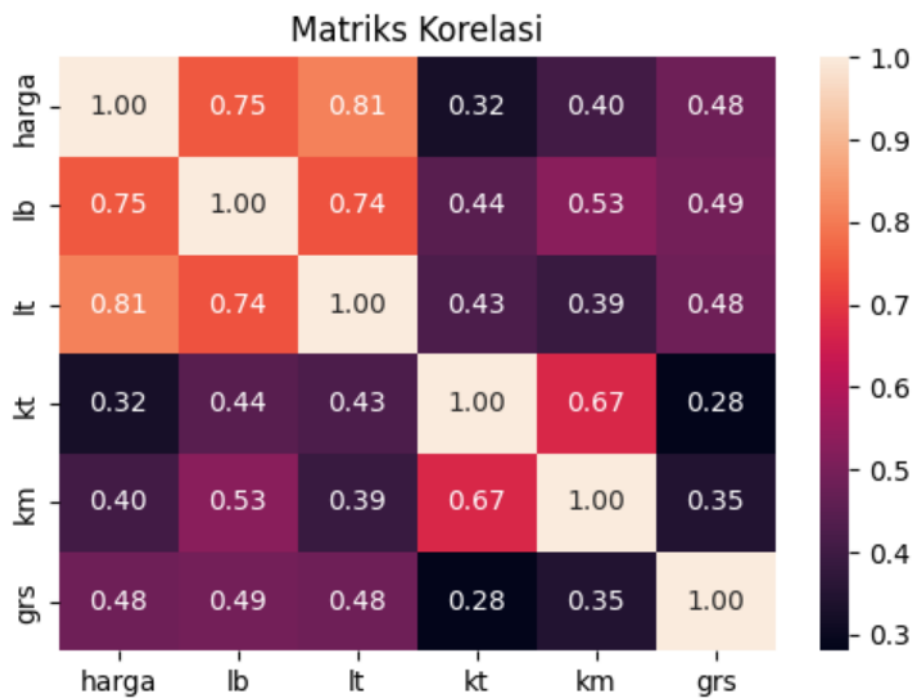
Mean Absolute Error (MAE)

Hasil evaluasi menunjukkan bahwa regresi linear dapat memberikan prediksi yang cukup akurat terhadap harga rumah di wilayah tersebut.

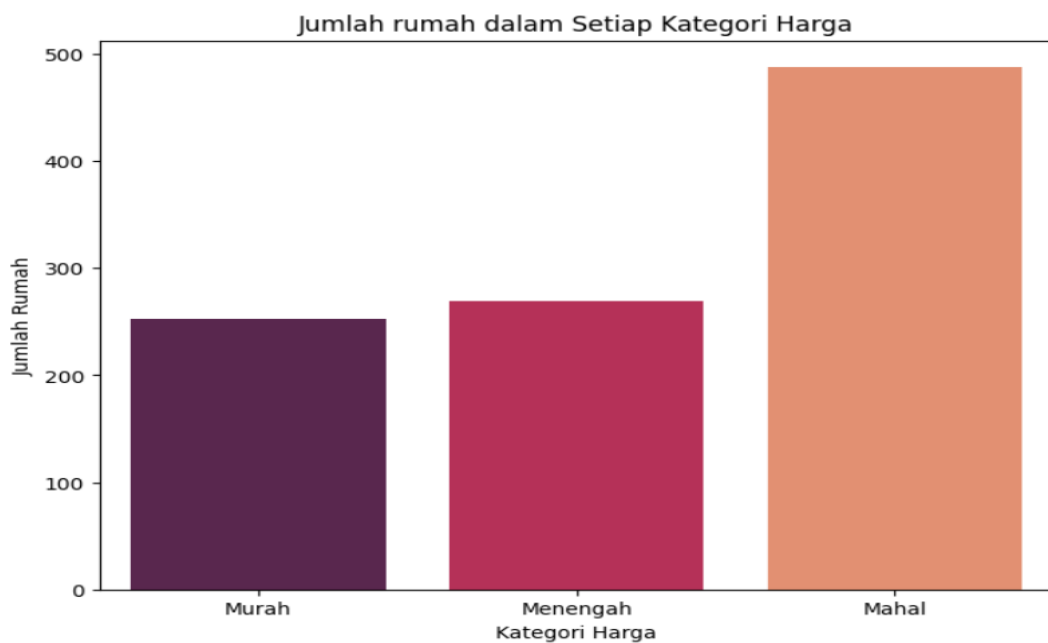
Visualisasi explorasi data



Matriks kolerasi



### Visualisasi tingkat harga



## Evaluasi model ridge

Evaluasi Model Ridge pada Data Pelatihan:

MAE: 2117.4069990688267

MSE: 16455645.413211435

R2 Score: 0.7040692834779283

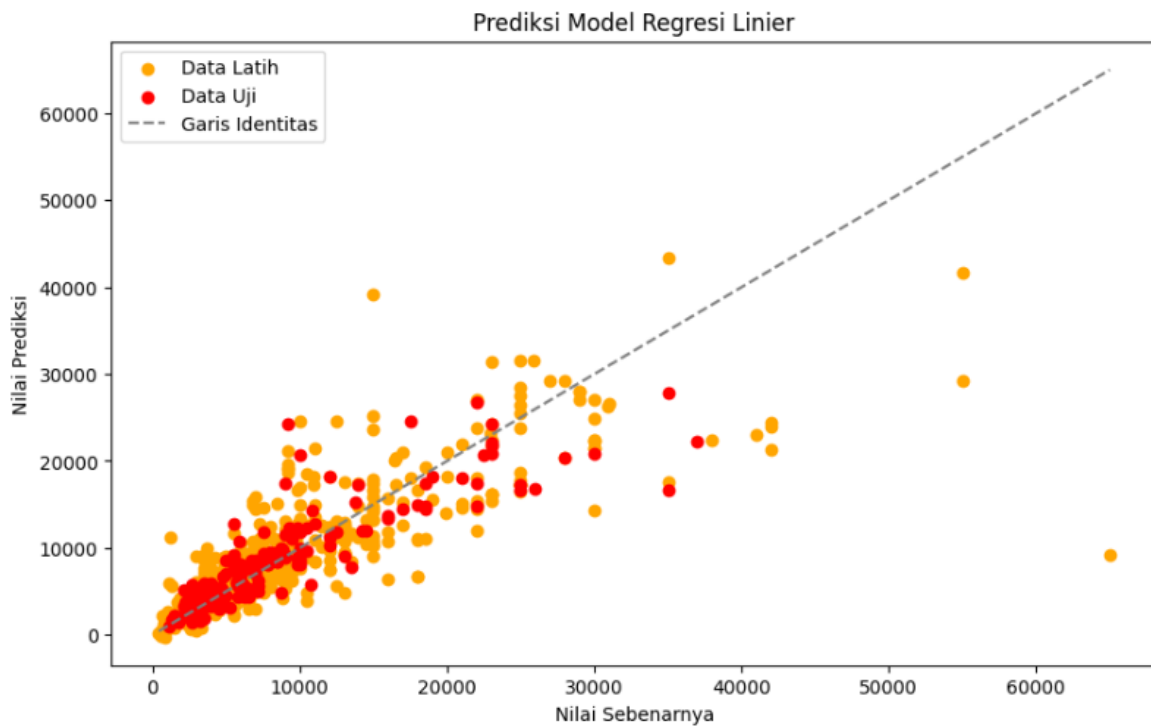
Evaluasi Model Ridge pada Data Pengujian:

MAE: 1980.3562527086858

MSE: 10675957.081648333

R2 Score: 0.7713084102181795

## Prediksi model regresi



## Tampilan hasil prediksi



Sesuaikan Rumah Kebutuhan Anda :

LB:  1120.00

LT:  1400.00

KT:  10.00

KM:  10.00

GRS:  10.00

Harga rumah impian anda diperkirakan sekitar IDR 47,217.437 juta