**PREPROCESSING REFINING PADA DATA REAL ESTATE UNTUK PREDIKSI HARGA RUMAH**

**PRAKTIKUM PEMROSESAN DATA MODUL 1 PART 2**

**Oleh**

**Nama**

**NIM**

**(Program Studi Informatika)**

**A black background with a black square

Description automatically generated with medium confidence**

**UNIVERSITAS TRISAKTI  
September 2024**

**I. Pendahuluan**

Pada praktikum ini akan dilakukan pembelajaran mengenai data real estate, preprocessing dan refining merupakan langkah-langkah yang sangat penting sebelum menerapkan model prediksi. Data real estate sering kali mencakup berbagai jenis informasi, seperti harga rumah, lokasi, ukuran, kondisi, dan fitur tambahan (misalnya, jumlah kamar tidur, taman, kolam renang). Data ini bisa sangat heterogen dan mengandung berbagai masalah seperti nilai yang hilang, outlier, dan variabel yang tidak relevan.

Preprocessing adalah proses membersihkan dan menyiapkan data untuk analisis lebih lanjut. Ini termasuk langkah-langkah seperti mengisi nilai yang hilang, menghapus duplikasi, normalisasi data, dan encoding variabel kategorikal. Refining, di sisi lain, melibatkan penyempurnaan dan transformasi data untuk meningkatkan kualitas dan relevansi data dalam konteks analisis. Ini mungkin termasuk teknik seperti pemilihan fitur, transformasi variabel, dan pembentukan variabel baru.

**Tujuan Praktikum**

Praktikum ini bertujuan untuk memberikan pemahaman mendalam mengenai teknik preprocessing dan refining data dalam konteks prediksi harga rumah. Dengan mempelajari langkah-langkah ini, peserta praktikum diharapkan dapat:

1. **Mengidentifikasi dan Mengatasi Masalah Data:** Memahami berbagai masalah yang mungkin muncul dalam data real estate dan cara mengatasinya melalui teknik preprocessing.
2. **Meningkatkan Kualitas Data:** Melalui refining, peserta akan belajar cara meningkatkan kualitas data untuk model prediksi, termasuk teknik pemilihan fitur dan transformasi variabel.
3. **Persiapan Data untuk Model Prediksi:** Menyiapkan data yang telah diproses dengan baik sehingga dapat digunakan dalam model prediksi harga rumah yang lebih akurat.

**II. Hasil Praktikum**

**II.1 *Import Library***

|  |
| --- |

| **Penjelasan:** |
| --- |

**II.2 Menyiapkan *Dataset***

|  |
| --- |

| **Penjelasan:** |
| --- |

**II.3 Melakukan Visualisasi variabel target dan Melakukan log transform**

|  |
| --- |

| **Penjelasan** |
| --- |

**Pertanyaan :**

**Lakukan uji normal dan visualisasi distribusi data pada variabel prediktor**

| **Screenshot :** |
| --- |

| **Penjelasan :** |
| --- |

**II.4 Melakukan uji hipotesa**

|  |
| --- |

| **Penjelasan :** |
| --- |

|  |
| --- |

| **Penjelasan :** |
| --- |

**II.5 Menggunakan log transform untuk mendekatkan distribusi data mendekati normal**

|  |
| --- |

| **Penjelasan:** |
| --- |

|  |
| --- |

| **Apakah sudah normal atau belum? Berikan alasannya!**  **Penjelasan :** |
| --- |

**II.6 Melakukan analisis dengan metode KMeans**

|  |
| --- |

| **Lakukan ujicoba untuk kandidat fitur yang cocok untuk dikelompokkan berdasarkan klaster dan temukan berapa nilai k optimal?**  **Penjelasan:** |
| --- |

**II.7 Melakukan clustering/ pengelompokan pada data dengan menggunakan KMeans**

|  |
| --- |

| **Penjelasan :** |
| --- |

**II.8 Kesimpulan ( minimal 5 baris )**

|  |
| --- |