Nama : Felicia Amanda Cahyadewi

NIM : A11.2022.14090

Kelompok : Penambangan Data – A11.4519

UTS

1. **Topik**

Analisis Segmentasi Pelanggan Mall Menggunakan K-Means Clustering Berdasarkan Pendapatan Tahunan dan Kebiasaan Belanja

1. **Deskripsi Singkat (Ringkasan BU/DU)**

Proyek ini berfokus pada pemahaman karakteristik dan perilaku belanja pelanggan di sebuah mall dengan tujuan akhir melakukan segmentasi yang lebih efektif. Segmentasi ini membantu pihak manajemen mall dalam menyusun strategi pemasaran yang lebih tepat sasaran bagi masing-masing segmen pelanggan. Tahapan proyek ini meliputi pemahaman data dan pemahaman tujuan bisnis untuk membangun model yang sesuai dengan kebutuhan.

1. **Masalah dan Tujuan yang akan diselesaikan**

Masalah yang diangkat dalam proyek ini adalah bagaimana mengidentifikasi karakteristik pelanggan yang memiliki kesamaan dalam profil dan perilaku belanja. Dengan adanya segmentasi ini, diharapkan mall dapat memberikan layanan yang lebih dipersonalisasi dan meningkatkan efektivitas kampanye pemasaran.

Tujuan akhir dari proyek ini adalah menghasilkan beberapa segmen pelanggan berdasarkan kesamaan fitur dan memberikan rekomendasi strategi pemasaran yang relevan untuk setiap segmen

1. **Penjelasan Datasets**

Dataset yang digunakan berasal dari data internal mall, berisi 200 entri mengenai pelanggan dengan fitur-fitur sebagai berikut:

Dataset : <https://www.kaggle.com/datasets/vjchoudhary7/customer-segmentation-tutorial-in-python>

**Penjelasan Atribut:**

* CustomerID: Nomor unik untuk setiap pelanggan.
* Gender: Jenis kelamin pelanggan, dengan nilai "Male" atau "Female".
* Age: Usia pelanggan, berupa angka integer yang menunjukkan usia dalam tahun.
* Annual Income (k$): Pendapatan tahunan pelanggan dalam ribuan dolar, digunakan untuk memahami status ekonomi pelanggan.
* Spending Score (1-100): Skor yang menunjukkan perilaku belanja pelanggan dalam skala 1 hingga 100; skor tinggi menunjukkan kecenderungan belanja yang lebih besar.

1. **Alur/Tahapan/Kerangka Eksperimen**

* Data Understanding: Eksplorasi awal pada dataset untuk memahami distribusi data dari setiap fitur, pengecekan kesalahan atau nilai yang hilang, dan analisis statistik deskriptif.
* Preprocessing Data: Melakukan pembersihan data, termasuk normalisasi pada fitur numerik agar setiap fitur memiliki skala yang sama sebelum diterapkan dalam proses clustering.
* Exploratory Data Analysis (EDA): Visualisasi data untuk menemukan pola atau korelasi antara fitur, seperti analisis hubungan antara pendapatan tahunan dan skor belanja.
* **Clustering dengan KMeans**: Segmentasi pelanggan dengan KMeans Clustering menggunakan jumlah cluster yang telah ditentukan, dan mengelompokkan pelanggan berdasarkan karakteristik yang mirip. Dengan menggunakan algoritma K-Means Clustering, pelanggan akan dikelompokkan berdasarkan variabel seperti usia, jenis kelamin, pendapatan tahunan, dan skor belanja.
* Evaluation: Mengevaluasi kualitas cluster dengan menggunakan silhouette score atau metrik lain yang sesuai untuk memastikan efektivitas segmentasi.
* Interpretasi Hasil: Menginterpretasikan hasil clustering dan mengidentifikasi karakteristik dari masingmasing segmen yang terbentuk.

1. **Timeline Eksperimen (sampai dengan akhir semester)**

|  |  |  |
| --- | --- | --- |
|  | **Minggu** | **Kegiatan** |
| 1 | Minggu ke-1 | **Pengumpulan data, pemahaman dataset**, dan analisis deskriptif awal. |
| 2 | Minggu ke-2 | **Exploratory Data Analysis** (EDA) untuk menemukan pola atau hubungan antar fitur. |
| 3 | Minggu ke-3 | **Preprocessing data**, normalisasi, dan pengecekan nilai yang hilang atau tidak sesuai. |
| 4 | Minggu ke-4 | **Preprocessing data**, normalisasi, dan pengecekan nilai yang hilang atau tidak sesuai. |
| 5 | Minggu ke-5 | Penerapan algoritma K-Means Clustering dan pembentukan cluster pelanggan. |
| 6 | Minggu ke-6 | **Evaluasi hasil** clustering untuk memastikan kualitas dan validitas segmen. |
| 7 | Minggu ke-7 | Interpretasi hasil clustering, karakteristik setiap segmen, dan penyusunan rekomendasi strategi pemasaran. |