**Analisis Segmentasi Pelanggan pada Data Transaksi Penjualan Restoran Cepat Saji Menggunakan Clustering dan Klasifikasi**



Disusun oleh:

Azzahro Hidayatus Sakdiyah

(A11.2022.14026)

A11.4504

**FAKULTAS ILMU KOMPUTER**

**UNIVERSITAS DIAN NUSWANTORO**

**SEMARANG**

**2024**

1. **Deskripsi Singkat**  
   Penelitian ini bertujuan untuk menganalisis data transaksi dari sebuah restoran cepat saji dan menerapkan metode clustering untuk segmentasi pelanggan serta klasifikasi untuk memprediksi perilaku pembelian. Dengan segmentasi yang dihasilkan, restoran dapat menyusun strategi pemasaran yang lebih tepat sasaran sesuai dengan kebutuhan dan karakteristik tiap segmen pelanggan.
2. **Masalah dan Tujuan yang akan Diselesaikan**  
   Masalah utama yang dihadapi adalah kurangnya pemahaman terhadap perilaku pelanggan dalam bertransaksi dan bagaimana menentukan strategi pemasaran yang efektif berdasarkan data tersebut. Penelitian ini bertujuan untuk:
   * Mengelompokkan pelanggan berdasarkan pola transaksi dan karakteristik pembelian dengan teknik clustering.
   * Memprediksi kecenderungan pembelian pelanggan pada kelompok tertentu berdasarkan klasifikasi.
   * Menyusun rekomendasi strategi pemasaran berbasis data untuk tiap segmen pelanggan yang terbentuk.
3. **Alur / Tahapan / Kerangka Eksperimen**
   1. **Data Preprocessing**  
      Membersihkan data, menangani nilai yang hilang, normalisasi data, dan encoding variabel kategorikal untuk memastikan data siap dianalisis.
   2. **Eksplorasi Data**  
      Melakukan eksplorasi data awal untuk memahami distribusi fitur, korelasi antar-fitur, dan pola-pola dasar yang mungkin ada.
   3. **Clustering untuk Segmentasi Pelanggan**  
      Menggunakan algoritma clustering, seperti K-Means atau DBSCAN, untuk mengelompokkan pelanggan berdasarkan fitur transaksi (misalnya, jumlah belanja, frekuensi kunjungan).
   4. **Klasifikasi untuk Memprediksi Perilaku Pembelian**  
      Menerapkan algoritma klasifikasi, seperti Decision Tree atau Random Forest, untuk memprediksi kecenderungan pembelian pada tiap segmen yang terbentuk. Model klasifikasi akan dilatih menggunakan data historis transaksi untuk memprediksi preferensi pelanggan.
   5. **Evaluasi Model dan Interpretasi Hasil**  
      Mengevaluasi kinerja model clustering dan klasifikasi berdasarkan metrik yang relevan, seperti silhouette score untuk clustering dan akurasi untuk klasifikasi. Hasilnya akan dianalisis untuk menyusun rekomendasi bagi manajemen restoran.
   6. **Penyusunan Laporan dan Kesimpulan**  
      Menyusun laporan akhir yang mencakup hasil segmentasi pelanggan, interpretasi dari masing-masing kelompok pelanggan, rekomendasi strategi pemasaran berbasis data, serta kesimpulan dari penelitian.
4. **Penjelasan Datasets**
   1. **Sumber Data:** Data diperoleh dari sistem transaksi restoran yang mencatat informasi penjualan selama periode tertentu.
   2. **Fitur-fitur dalam Dataset:**

i. CustomerID: ID unik untuk setiap pelanggan

ii. Age: Usia pelanggan

iii. Gender: Jenis kelamin pelanggan

iv. Income: Pendapatan pelanggan

v. VisitFrequency: Frekuensi kunjungan ke restoran

vi. AverageSpend: Rata-rata pengeluaran per kunjungan

vii. PreferredCuisine: Jenis masakan yang disukai

viii. TimeOfVisit: Waktu kunjungan (misalnya: makan siang, makan malam)

ix. GroupSize: Jumlah orang dalam satu kunjungan

x. DiningOccasion: Ocasion kunjungan (misalnya: casual, bisnis)

xi. MealType: Jenis makanan yang dipesan

xii. OnlineReservation: Apakah melakukan reservasi online

xiii. DeliveryOrder: Apakah menggunakan layanan pesan antar

xiv. LoyaltyProgramMember: Status keanggotaan program loyalitas

xv. WaitTime: Waktu tunggu pelayanan

xvi. ServiceRating: Penilaian terhadap layanan

xvii. FoodRating: Penilaian terhadap makanan

xviii. AmbianceRating: Penilaian terhadap suasana restoran

xix. HighSatisfaction: Target variabel yang menunjukkan tingkat kepuasan tinggi

* 1. **Exploratory Data Analysis (EDA)**
* **Analisis Deskriptif:** Rata-rata jumlah pembelian, distribusi nilai transaksi, pola kunjungan.
* **Visualisasi:** Histogram, scatter plot, heatmap korelasi antar fitur.
* **Outlier Detection:** Identifikasi data yang tidak wajar atau ekstrim.
  1. **Proses Feature Dataset**
* **Pembersihan Data:** Menghapus data duplikat dan menangani nilai yang hilang.
* **Normalisasi:** Skala fitur menggunakan Min-Max Scaling atau Standarisasi.
* **Feature Engineering:** Membuat fitur baru seperti RFM (Recency, Frequency, Monetary).

1. **Proses Learning/Modeling**

* **Clustering:**
* Algoritma : K-Means.
* Evaluasi : Elbow Method untuk menentukan jumlah klaster optimal.
* Output : Label klaster untuk setiap pelanggan.
* **Klasifikasi:**
* Algoritma : Random Forest, Decision Tree, atau Logistic Regression.
* Validasi : Cross-validation dan metrics (Accuracy, F1-Score, Precision, Recall).

1. **Performa Model**

* **Clustering:**
* Silhouette Score: Mengukur kualitas klaster.
* Pemisahan yang jelas antar klaster di visualisasi (PCA/TSNE).
* **Klasifikasi :**
  + Confusion Matrix: Analisis prediksi benar/salah.
  + ROC Curve: Evaluasi trade-off antara True Positive Rate dan False Positive Rate.

1. **Diskusi Hasil dan Kesimpulan**

* **Diskusi Hasil:**
* Segmen pelanggan dengan karakteristik tertentu (misalnya, pelanggan dengan frekuensi tinggi namun nilai pembelian rendah).
* Efektivitas model klasifikasi dalam memprediksi segmen pelanggan baru.
* **Kesimpulan:**
* Segmentasi pelanggan memberikan wawasan mendalam tentang perilaku pelanggan.
* Rekomendasi untuk meningkatkan strategi pemasaran seperti promosi personal atau program loyalitas.

Dengan pendekatan ini, diharapkan restoran dapat memahami pelanggan mereka lebih baik dan merancang kampanye pemasaran yang lebih tepat sasaran serta meningkatkan kepuasan pelanggan.