

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Perkembangan teknologi digital telah mendorong perubahan besar dalam pola konsumsi masyarakat, termasuk dalam layanan pemesanan makanan secara daring. Go Food menjadi salah satu platform yang paling banyak digunakan karena menawarkan kemudahan, kecepatan, serta variasi menu yang semakin beragam [1]. Pertumbuhan ini menyebabkan persaingan antar merchant meningkat, khususnya dalam aspek harga dan diskon yang ditampilkan kepada konsumen [2]. Banyaknya pilihan menu yang muncul secara bersamaan dalam aplikasi membuat konsumen kesulitan membedakan rentang harga suatu menu, sementara merchant menghadapi tantangan dalam menampilkan menu secara strategis agar mampu bersaing secara optimal.

Permasalahan utama muncul ketika menu-menu Go Food tidak memiliki struktur pengelompokan yang jelas berdasarkan harga dan diskon. Banyak merchant masih menerapkan strategi harga secara manual tanpa analisis berbasis data, sehingga menu yang ditampilkan dalam aplikasi menjadi tidak terstruktur dan kurang relevan dengan preferensi harga pengguna [3]. Kondisi ini dapat menurunkan efektivitas promosi, mengurangi tingkat konversi pembelian, dan berpotensi menimbulkan kerugian bagi merchant. Ketidakteraturan informasi harga tersebut memperbesar tantangan bagi konsumen dalam membandingkan opsi menu, khususnya pada platform dengan jumlah produk yang sangat besar seperti Go Food.

Dari penelitian sebelumnya menerapkan K-Means Clustering dalam konteks segmentasi pelanggan, analisis perilaku konsumen, dan pemetaan produk, masih sangat sedikit [4]. Penelitian yang secara khusus mengelompokkan menu makanan berdasarkan kombinasi harga dan diskon pada layanan pesan-antar. Padahal kedua atribut tersebut merupakan indikator utama dalam persaingan merchant yang kompetitif. Ketersediaan dataset harga dan diskon di platform Go Food menunjukkan peluang besar untuk melakukan segmentasi berbasis data, sehingga kesenjangan penelitian yang perlu diisi melalui pendekatan yang sesuai.

Untuk menjawab permasalahan tersebut diperlukan metode analisis yang efisien, cepat, dan mampu mengelompokkan data numerik secara jelas. Algoritma K-Means Clustering relevan digunakan karena dirancang untuk menangani data numerik kontinue seperti price dan discount\_price, memiliki kinerja yang baik pada dataset berukuran besar, serta menghasilkan kluster yang mudah diinterpretasikan [5]. Penelitian terdahulu juga menunjukkan bahwa K-Means memiliki performa tinggi dalam segmentasi berbasis jarak ketika metrik yang tepat digunakan [6]. Dengan demikian, penggunaan K-Means dalam penelitian ini diharapkan mampu menghasilkan pengelompokan menu Go Food berdasarkan harga dan diskon yang lebih terstruktur serta dapat dimanfaatkan untuk mendukung strategi harga dan promosi berbasis data.

## **1.2 Rumusan Masalah**

Berdasarkan latar belakang yang telah diuraikan, serta mempertimbangkan keterbatasan data yang digunakan dari sumber publik (Kaggle), maka rumusan masalah dalam penelitian ini dapat dirumuskan sebagai berikut:

1. Bagaimana penerapan algoritma K-Means Clustering dalam mengelompokkan produk makanan GoFood berdasarkan atribut harga dan harga diskon pada dataset publik dari Kaggle?
2. Berapa jumlah cluster yang optimal berdasarkan hasil evaluasi menggunakan metode Elbow dan Silhouette Score?
3. Apa karakteristik masing-masing kluster yang terbentuk dalam distribusi harga dan diskon, serta bagaimana interpretasi hasil tersebut dapat digunakan untuk mendukung strategi promosi produk?

### 1.3 Batasan Masalah

Berdasarkan perumusan masalah, peneliti memberikan batasan masalah sehingga tidak terjadi pembahasan diluar pokok bahasan yang dimaksud. Adapun batasan masalah sebagai berikut:

1. Penelitian ini menggunakan data menu aplikasi pemesanan makanan online (Go Food) yang tersedia dari sumber terbuka seperti dari platform Kaggle. Data yang diambil mencakup informasi terkait harga (*price*) dan harga diskon (*discount\_price*) dari berbagai menu.
2. Pengelompokan menu hanya dilakukan berdasarkan dua fitur numerik, yaitu harga dan harga diskon. Fitur-fitur lain yang tersedia dalam dataset (seperti nama merchant, area, kategori, display, nama produk) digunakan untuk analisis deskriptif dan interpretasi hasil clustering, tetapi tidak dimasukkan langsung sebagai fitur dalam proses clustering K-Means.
3. Algoritma pengelompokan yang digunakan dalam penelitian ini hanya K-Means Clustering. Penelitian ini tidak membandingkan atau menggunakan algoritma clustering lainnya.
4. Penentuan jumlah cluster optimal utama didasarkan pada Metode Elbow dan diinterpretasikan berdasarkan karakteristik cluster yang terbentuk. Metode penentuan jumlah cluster lainnya (misalnya, Silhouette Analysis secara ekstensif, Davies-Bouldin Index sebagai penentu utama) tidak menjadi fokus utama dalam menentukan jumlah cluster yang digunakan.
5. Interpretasi cluster dilakukan berdasarkan rata-rata harga, diskon, dan distribusi fitur kategorikal dalam setiap cluster. Analisis mendalam terkait faktor-faktor lain yang mungkin mempengaruhi pengelompokan (selain harga dan diskon) tidak dibahas secara ekstensif di luar data yang tersedia.

#### 1.4 Tujuan Penelitian

Adapun Tujuan dari dibuatnya penelitian ini adalah sebagai berikut:

1. Menerapkan algoritma K-Means Clustering untuk mengelompokkan produk makanan dalam platform Go Food berdasarkan dua atribut numerik, yaitu harga dan harga diskon, menggunakan dataset publik dari Kaggle.
2. Menentukan jumlah cluster yang optimal melalui evaluasi model menggunakan metode Elbow dan Silhouette Score.
3. Menganalisis dan menginterpretasikan karakteristik masing-masing kluster berdasarkan distribusi nilai harga dan diskon, serta mengevaluasi potensi strategi promosi yang sesuai dengan masing-masing segmen produk.

#### 1.5 Manfaat Penelitian

Adapun manfaat dari dibuatnya penelitian ini adalah sebagai berikut:

##### 1. Manfaat Teoritis

Penelitian ini diharapkan dapat memberikan kontribusi terhadap pengembangan ilmu pengetahuan, khususnya dalam bidang data mining dan analisis kluster berbasis algoritma K-Means. Hasil dari penelitian ini dapat memperkaya literatur mengenai penerapan metode unsupervised learning dalam segmentasi produk makanan berdasarkan parameter numerik seperti harga dan diskon. Selain itu, penelitian ini juga dapat menjadi referensi bagi penelitian selanjutnya yang ingin mengkaji penerapan K-Means dalam konteks promosi digital, e-commerce, atau sistem rekomendasi berbasis segmentasi.

## 2. Manfaat Praktis

Secara praktis, hasil penelitian ini dapat dimanfaatkan oleh pelaku usaha kuliner, khususnya merchant pada platform online food delivery seperti GoFood, untuk memahami struktur segmentasi produk berdasarkan harga dan diskon. Dengan informasi ini, merchant dapat menyusun strategi penetapan harga dan promosi yang lebih tepat sasaran sesuai dengan karakteristik setiap kluster produk. Selain itu, penelitian ini juga dapat memberikan insight awal bagi pengembang sistem atau analis data dalam merancang solusi pemasaran berbasis data yang bersifat adaptif terhadap perilaku konsumen dan kompetisi pasar.

### 1.6 Sistematika Penulisan

Sistematika penulisan dan urutan tata letak yang ada dalam tugas akhir ini adalah sebagai berikut:

**BAB I PENDAHULUAN**, Bab ini berfungsi sebagai pengantar penelitian, memaparkan konteks yang relevan mengenai perkembangan layanan pesan-antar makanan seperti Go Food dan pentingnya analisis data dalam industri ini, khususnya terkait strategi harga dan diskon menu. Bab ini juga merumuskan permasalahan yang spesifik, menetapkan tujuan yang ingin dicapai melalui analisis pengelompokan menu menggunakan K-Means Clustering.

**BAB II TINJAUAN PUSTAKA**, Bab ini menyajikan landasan teoritis yang mendukung penelitian. Pembahasan dimulai dengan konsep-konsep dasar layanan pesan-antar makanan dan pentingnya strategi harga serta diskon dalam konteks bisnis kuliner digital. Kemudian, bab ini mendalami teori mengenai data mining, konsep clustering secara umum, dan secara spesifik menjelaskan algoritma K-Means Clustering beserta mekanisme kerjanya. Metode-metode evaluasi hasil clustering seperti Metode Elbow, Silhouette Coefficient, dan Davies-Bouldin Index juga dibahas sebagai alat ukur kualitas pengelompokan.

**BAB III METODOLOGI PENELITIAN**, Bab ini menguraikan secara rinci pendekatan dan prosedur yang ditempuh dalam melaksanakan penelitian. Dimulai dengan penjelasan mengenai jenis dan pendekatan penelitian, bab ini kemudian mendeskripsikan secara lengkap sumber data yang digunakan, karakteristiknya, dan bagaimana data tersebut dikumpulkan. Variabel-variabel yang terlibat dalam analisis, terutama fitur harga dan harga diskon untuk clustering, didefinisikan. Bagian utama bab ini adalah penjelasan tahapan penelitian secara sistematis, mencakup proses pemahaman data, pembersihan data, analisis deskriptif (EDA), persiapan data untuk clustering (termasuk penskalaan), penerapan algoritma K-Means (bersama dengan penentuan jumlah cluster), evaluasi hasil, hingga proses interpretasi cluster. Metode analisis data yang spesifik untuk setiap tahapan juga dijelaskan.

**BAB IV HASIL DAN PEMBAHASAN**, Bab ini menyajikan temuan dari analisis data. Diawali dengan deskripsi data setelah proses pembersihan dan menampilkan hasil analisis deskriptif awal (EDA) berupa ringkasan statistik dan visualisasi distribusi fitur. Hasil penentuan jumlah cluster optimal menggunakan Metode Elbow disajikan, di ikuti dengan visualisasi hasil penerapan K-Means Clustering yang menunjukkan pengelompokan menu berdasarkan harga dan diskon. Kualitas clustering dievaluasi dan dilaporkan menggunakan metrik Silhouette Coefficient dan Davies-Bouldin Index.

**BAB V PENUTUP**, Bab terakhir ini menyimpulkan seluruh rangkaian penelitian. Kesimpulan menyajikan jawaban ringkas atas rumusan masalah dan merangkum karakteristik utama dari cluster-cluster menu yang berhasil diidentifikasi. Bab ini juga mengidentifikasi keterbatasan yang dihadapi selama pelaksanaan penelitian. Terakhir, bab ini memberikan saran-saran konstruktif untuk penelitian di masa mendatang, mengusulkan arah pengembangan atau eksplorasi lebih lanjut terkait topik analisis data menu Go Food atau penerapan teknik clustering lainnya.