

BAB I

PENDAHULUAN

1.1 Latar Belakang

Dalam era digital yang sedang berlangsung, dimana teknologi informasi terus berkembang dengan pesat, sektor e-commerce telah menjadi salah satu industri yang mengalami pertumbuhan luar biasa. Di Indonesia, platform e-commerce terkemuka seperti Tokopedia telah berhasil mencatat prestasi sebagai pemain utama dalam pasar ini. Tokopedia menyediakan berbagai macam produk, termasuk laptop, dengan beragam merek, spesifikasi, dan harga. Namun, dengan banyaknya pilihan yang tersedia di platform ini, konsumen sering kali menghadapi tantangan dalam memilih produk yang paling sesuai dengan kebutuhan dan preferensi mereka. Oleh karena itu, diperlukan pendekatan yang sangat efektif untuk menganalisis data produk laptop yang ada di Tokopedia, dengan tujuan memberikan rekomendasi yang lebih akurat dan personal kepada konsumen.

Salah satu metode yang umum digunakan dalam analisis data untuk mengatasi tantangan ini adalah metode clustering. Metode ini digunakan untuk mengelompokkan objek-objek serupa berdasarkan kesamaan karakteristik tertentu. Dalam konteks analisis data produk laptop di Tokopedia, metode clustering dapat sangat membantu dalam mengidentifikasi kelompok produk yang memiliki fitur serupa, seperti merek, jenis prosesor, ukuran layar, dan harga. Dengan menggunakan teknik clustering, kita dapat mengelompokkan berbagai produk laptop ke dalam beberapa kluster yang berbeda, sehingga mempermudah proses memberikan rekomendasi produk yang relevan kepada konsumen.

Namun, meskipun sudah ada penelitian sebelumnya yang mencoba menerapkan metode clustering pada data produk di platform e-commerce, seringkali penelitian tersebut terbatas pada penggunaan metode tertentu dan kurang dalam hal membandingkan kinerja antara berbagai metode clustering yang berbeda. Oleh karena itu, penelitian yang lebih komprehensif diperlukan untuk mengisi celah ini. Penelitian ini bertujuan untuk melakukan studi perbandingan antara tiga metode clustering yang umum digunakan, yaitu K-Means, Hierarchical, dan DBSCAN.

Dalam rangka mengisi kekosongan ini, penelitian ini akan memperhatikan kinerja ketiga metode clustering ini dalam konteks analisis data produk laptop di Tokopedia. Penelitian ini akan melibatkan penggunaan berbagai ukuran evaluasi clustering yang telah diakui dalam literatur, seperti Silhouette Index, Indeks Calinski-Harabasz, dan Indeks Davies-Bouldin. Dengan membandingkan hasil dari ketiga metode clustering ini, kita akan dapat memahami kelebihan dan kekurangan masing-masing metode dalam menghadapi tantangan analisis data produk laptop di Tokopedia.

Hasil yang diharapkan dari penelitian ini akan memiliki dampak penting dalam mengoptimalkan sistem rekomendasi produk di platform e-

commerce seperti Tokopedia. Selain itu, penelitian ini juga diharapkan dapat meningkatkan pengalaman belanja online konsumen, dengan memberikan rekomendasi yang lebih relevan dan personal berdasarkan karakteristik produk yang sesuai dengan kebutuhan dan preferensi mereka. Dengan demikian, penelitian ini bukan hanya akan memberikan wawasan baru dalam analisis data e-commerce, tetapi juga dapat membawa dampak positif dalam dunia nyata bagi konsumen dan penjual di platform e-commerce.

Setelah berhasil membandingkan ketiga metode *clustering*. Tahap selanjutnya adalah membuat rekomendasi dengan membandingkan dua metode rekomendasi yaitu *Content Based Filtering* dan *Item Based Collaborative Filtering*.

1.2 Perumusan Masalah

Dalam konteks perkembangan e-commerce dan kemajuan teknologi informasi, penelitian ini diilhami oleh kebutuhan untuk meningkatkan pengalaman belanja online konsumen di platform e-commerce, terutama di Tokopedia, yang merupakan salah satu platform terbesar di Indonesia. Meskipun berbagai produk laptop dengan beragam merek, spesifikasi, dan harga ditawarkan, konsumen sering menghadapi kesulitan dalam memilih produk yang paling sesuai dengan kebutuhan mereka. Oleh karena itu, penelitian ini akan memusatkan perhatian pada penerapan metode clustering dalam analisis data produk laptop di Tokopedia, dengan tujuan untuk memberikan rekomendasi yang lebih akurat dan personal kepada konsumen.

Namun, penelitian-penelitian sebelumnya dalam domain ini masih memiliki beberapa keterbatasan. Terutama, fokus penelitian sebelumnya sering terbatas pada penerapan metode clustering tertentu tanpa membandingkan kinerja dengan metode clustering alternatif. Oleh karena itu, rumusan masalah yang relevan dengan latar belakang di atas adalah:

1. Bagaimana penerapan metode clustering, seperti K-Means, Hierarchical, dan DBSCAN, dalam analisis data produk laptop di Tokopedia dapat membantu mengelompokkan produk-produk serupa berdasarkan karakteristik tertentu?
2. Bagaimana penggunaan ukuran evaluasi clustering seperti Silhouette Index, Indeks Calinski-Harabasz, dan Indeks Davies-Bouldin dapat membantu dalam membandingkan kinerja metode clustering yang berbeda?
3. Bagaimana perbandingan kinerja dari dua algoritma rekomendasi dalam memberikan rekomendasi produk berdasarkan kriteria atau karakteristik item yang diinginkan?

1.3 Tujuan Penelitian

Penelitian ini bertujuan untuk menerapkan tiga metode clustering (K-Means, Hierarchical dan DBSCAN) pada kasus analisa data produk laptop di E-Commerce Tokopedia. Dan setelah itu akan membandingkan kinerja dari tiga

algoritma clustering yang cukup populer digunakan tersebut. Kemudian membuat sistem rekomendasi menggunakan dua algoritma rekomendasi yaitu *Content Based Filtering* dan *Item-Based Collaborative Filtering*.

1.4 Batasan Masalah

Batasan masalah dalam penelitian ini penting untuk mengarahkan fokus penelitian dan membatasi ruang lingkup agar tetap dapat dijalankan secara efisien. Berdasarkan latar belakang yang telah dijelaskan, berikut adalah beberapa batasan masalah yang dapat diidentifikasi:

1. Data dan Platform: Penelitian ini akan berfokus pada analisis data produk laptop yang tersedia di platform e-commerce Tokopedia.
2. Metode Clustering: Penelitian ini akan membandingkan tiga metode clustering yang umum digunakan, yaitu K-Means, Hierarchical, dan DBSCAN.
3. Variabel Analisis: Variabel yang akan menjadi fokus analisis dalam penelitian ini adalah merek laptop, jenis prosesor, harga, ukuran storage dan RAM.
4. Ukuran Evaluasi: Penelitian ini akan menggunakan tiga ukuran evaluasi clustering yang umum digunakan, yaitu Silhouette, Calinski-Harabasz, dan Davies-Bouldin.
5. Waktu Data: Penelitian ini akan dilakukan dengan menggunakan data yang tersedia sampai dengan batas waktu tertentu, yaitu tanggal 16 – 17 agustus 2023. Data yang lebih baru tidak akan dimasukkan dalam penelitian ini.

1.5 Manfaat Penelitian

Penelitian ini diharapkan akan memberikan beberapa manfaat, antara lain:

1. Memberikan wawasan lebih dalam tentang penerapan metode clustering dalam analisis data produk laptop di Tokopedia. Hal ini dapat membantu para pemangku kepentingan dalam memahami cara-cara yang efektif untuk mengelompokkan dan menganalisis data produk dalam konteks e-commerce.
2. Menawarkan perbandingan kinerja antara tiga metode clustering yang umum digunakan (K-Means, Hierarchical, dan DBSCAN) pada hasil skor matriks Silhouette, Calinski-Harabasz dan Davies-Bouldin. Hasil perbandingan ini dapat membantu pemilihan metode clustering yang paling sesuai dalam analisis data produk.
3. Menghasilkan sistem rekomendasi produk yang akurat dan personalisasi bagi pengguna. Dengan membandingkan dua algoritma rekomendasi, akan mendapatkan kesimpulan algoritma mana yang lebih baik untuk memberikan rekomendasi produk dengan menggunakan data yang dimiliki berdasarkan kesamaan karakteristik.

4. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi yang berarti dalam meningkatkan pemahaman dan penerapan metode clustering dalam analisis data produk, serta meningkatkan kualitas rekomendasi produk di industri *e-commerce*.