

BAB I

PENDAHULUAN

1.1 Latar Belakang

Sistem rekomendasi produk telah menjadi fitur penting dalam e-commerce untuk meningkatkan kepuasan pengguna dan mendorong konversi penjualan. Salah satu pendekatan paling populer dalam sistem ini adalah Collaborative Filtering (CF), yang memanfaatkan interaksi historis pengguna, seperti riwayat pembelian atau rating produk, untuk memprediksi preferensi pengguna lain [1][2].

Dalam pengembangan Collaborative Filtering, terdapat dua metode utama yang sering dibandingkan, yaitu Singular Value Decomposition (SVD) dan K-Nearest Neighbors (KNN) [3]. SVD bekerja dengan memetakan data interaksi pengguna dan produk ke dalam ruang dimensi laten melalui dekomposisi matriks, yang memungkinkan sistem menangkap pola tersembunyi meskipun data bersifat sparse. Sebaliknya, KNN menggunakan pendekatan berbasis neighborhood, dengan mencari entitas (pengguna atau produk) yang paling mirip untuk dijadikan referensi rekomendasi [4].

Kelebihan utama dari SVD adalah akurasinya yang tinggi dalam menangani data yang besar dan sparsity tinggi [5], sedangkan KNN unggul dalam hal kemudahan implementasi dan interpretasi, namun mengalami penurunan performa dalam skenario data berskala besar [6].

Di tengah pesatnya perkembangan e-commerce, pengguna kini dihadapkan pada jutaan produk, yang justru menimbulkan information overload. Sistem rekomendasi berperan menyaring informasi ini dan menyajikan produk yang relevan berdasarkan preferensi pengguna, sehingga memberikan pengalaman yang lebih personal dan efisien [7].

Meskipun beberapa studi menemukan bahwa SVD umumnya lebih unggul dalam hal akurasi dan skalabilitas dibanding KNN [3][5][8], belum ada kesimpulan mutlak mengenai model mana yang paling efektif dalam setiap konteks. Oleh karena itu, diperlukan evaluasi komparatif lebih lanjut untuk menguji performa

kedua model dalam situasi nyata, khususnya pada data e-commerce yang besar dan dinamis.

Sebagai respons terhadap kebutuhan tersebut, penelitian ini akan melakukan perbandingan menyeluruh antara model SVD dan KNN pada sistem rekomendasi produk berbasis Collaborative Filtering, menggunakan dataset perilaku konsumen supermarket daring. Perbandingan akan mencakup aspek akurasi prediksi (seperti RMSE dan Precision@K), efisiensi komputasi (waktu eksekusi dan penggunaan memori), serta sensitivitas terhadap fitur-fitur perilaku dan temporal pengguna. Hasilnya diharapkan dapat memberikan kontribusi praktis dan akademik dalam pemilihan algoritma sistem rekomendasi yang sesuai dengan kebutuhan dan karakteristik data.

1.2 Rumusan Masalah

1. Bagaimana hasil performa dan efektivitas model Singular Value Decomposition (SVD) dan K-Nearest Neighbors (KNN) dalam sistem rekomendasi produk berbasis Collaborative Filtering pada dataset besar dan sparse Hunter's Supermarket yang digunakan?
2. Bagaimana hasil perbandingan model yang lebih unggul dalam menghasilkan rekomendasi produk yang akurat dan relevan, serta efisien secara komputasi, ketika diaplikasikan pada data perilaku konsumen supermarket?
3. Bagaimana cara mengetahui faktor yang mempengaruhi performa masing-masing model (SVD dan KNN) dalam memberikan rekomendasi pada sistem e-commerce, khususnya berdasarkan fitur perilaku dan temporal pengguna?

1.3 Batasan Masalah

Penelitian ini memiliki sejumlah batasan yang disusun sebagai berikut:

1. Fokus penelitian hanya pada pendekatan Collaborative Filtering, dengan membandingkan dua model utama, yaitu SVD dan KNN.
2. Data yang digunakan merupakan dataset perilaku konsumen supermarket yang berisi transaksi pembelian dengan atribut seperti

- user_id, product_id, order_dow, dan reordered bernama Hunter's Supermarket dari Kaggle, tanpa pengumpulan data eksternal tambahan.
3. Evaluasi kinerja model akan menggunakan metrik akurasi seperti RMSE dan Recall, serta mempertimbangkan efisiensi komputasi dalam proses rekomendasi.
 4. Penelitian tidak membahas aspek bisnis secara mendalam, melainkan fokus pada aspek teknis dan performa algoritma rekomendasi.

1.4 Tujuan dan Manfaat Penelitian

1.4.1 Tujuan Penelitian

1. Menganalisis dan mengevaluasi performa serta efektivitas model Singular Value Decomposition (SVD) dan K-Nearest Neighbors (KNN) dalam sistem rekomendasi produk berbasis Collaborative Filtering pada dataset besar dan sparsity tinggi seperti Hunter's Supermarket.
2. Membandingkan akurasi, relevansi hasil rekomendasi, dan efisiensi komputasi dari model SVD dan KNN untuk menentukan model yang paling unggul dalam konteks aplikasi e-commerce.
3. Mengidentifikasi dan mengevaluasi faktor-faktor yang mempengaruhi performa masing-masing model, terutama fitur perilaku pengguna (seperti reordered) dan fitur temporal (seperti order_dow, days_since_prior_order), dalam proses pemberian rekomendasi.

1.4.2 Manfaat Penelitian

Penelitian ini diharapkan memberikan kontribusi jangka panjang dalam pengembangan sistem rekomendasi, khususnya dalam pemilihan algoritma yang tepat untuk data berskala besar dan sparsity tinggi. Dengan menyediakan analisis komparatif antara model SVD dan KNN, penelitian ini menjadi landasan bagi pengembangan metode yang lebih akurat, efisien, dan adaptif terhadap perubahan karakteristik data. Selain itu, hasil penelitian ini dapat memperkaya kajian ilmiah di bidang sistem rekomendasi berbasis Collaborative Filtering dan mendorong penelitian lanjutan yang lebih mendalam, termasuk eksplorasi model hybrid dan penerapan dalam konteks data nyata yang semakin kompleks.

1.5 Sistematika Penulisan

Berisi penjelasan singkat struktur isi mulai dari pendahuluan hingga kesimpulan dan proposal pada saat menulis laporan penelitian. Berikut cara menulis laporannya:

1. Bab 1 Pendahuluan:

Berisi uraian mengenai latar belakang permasalahan yang mendasari penelitian, perumusan masalah yang akan dijawab, batasan ruang lingkup penelitian, tujuan yang ingin dicapai, manfaat yang diharapkan dari hasil penelitian, serta penjelasan sistematika penulisan laporan secara keseluruhan.

2. Bab 2 Landasan Teori:

Menguraikan teori dasar mengenai sistem rekomendasi produk, Collaborative Filtering, model SVD dan KNN, serta studi literatur terkait.

3. Bab 3 Metodologi Penelitian:

Menjelaskan metode penelitian, deskripsi dataset, proses pra-pemrosesan data, serta implementasi dan evaluasi model.

4. Bab 4 Hasil dan Pembahasan:

Menyajikan hasil eksperimen, analisis perbandingan performa model, dan pembahasan temuan penelitian.

5. Bab 5 Simpulan dan Saran:

Menarik kesimpulan dan memberikan rekomendasi untuk penelitian selanjutnya atau implementasi praktis.