

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Di era transformasi digital yang berkembang pesat saat ini, teknologi internet telah secara fundamental mengubah pola konsumsi hiburan masyarakat global. Industri perfilman dan layanan streaming video (*Video on Demand*) mengalami pertumbuhan eksponensial, menyediakan akses instan terhadap katalog konten yang sangat masif. Platform seperti *Netflix*, *Disney+*, hingga layanan lokal menawarkan ribuan hingga jutaan judul film yang dapat diakses kapan saja. Namun, kelimpahan pilihan ini memunculkan paradoks psikologis yang dikenal sebagai *information overload*. Menurut penelitian yang dilakukan oleh Fayyaz, et al [1] menjelaskan bahwa fenomena ini terjadi ketika volume informasi yang tersedia melampaui kapasitas pemrosesan kognitif pengguna, menyebabkan kebingungan dan kesulitan dalam mengambil keputusan. Akibatnya, pengguna sering menghabiskan waktu yang tidak proporsional hanya untuk memilih tontonan, yang pada akhirnya dapat menurunkan tingkat kepuasan terhadap layanan penyedia konten tersebut.

Menyadari tantangan *information overload* tersebut, penyedia layanan digital kini dihadapkan pada urgensi bisnis yang sangat kritis. Bukti permasalahan ini secara nyata disoroti dalam penelitian Munthe & Siregar [2], yang menekankan bahwa kegagalan platform dalam menyajikan konten yang relevan berbanding lurus dengan penurunan tingkat retensi pengguna. Pengguna modern memiliki toleransi waktu yang sangat singkat; jika mereka terus-menerus kesulitan menavigasi tumpukan informasi yang masif, mereka cenderung akan meninggalkan layanan tersebut (*churn rate* meningkat). Oleh karena itu, Sistem Rekomendasi (*Recommender System*) berevolusi dari sekadar fitur tambahan menjadi instrumen strategis penentu keberlangsungan platform. Hal ini menuntut adanya intervensi komputasional adaptif yang mampu memetakan pola selera pengguna secara otomatis guna memberikan arahan tontonan yang tepat sasaran [3].

Secara taksonomi, upaya untuk memecahkan permasalahan komputasional tersebut bermuara pada tiga pendekatan algoritma yang paling fundamental: *Demographic Filtering*, *Content-Based Filtering*, dan *Collaborative Filtering*. Namun, dalam survei terbarunya, Li et al. [4] menyoroti sebuah permasalahan krusial di tingkat teknis: ketiga pendekatan ini memiliki filosofi yang sangat bertolak belakang dalam mendefinisikan kata "relevan". *Demographic* cenderung meratakan selera pengguna secara umum, *Content-Based* rentan menjebak pengguna dalam pilihan yang monoton, sedangkan *Collaborative* sangat bergantung pada kualitas riwayat data. Mengingat kompleksitas masalah tersebut, membandingkan ketiga metode ini secara empiris menjadi sebuah keharusan. Hal ini penting untuk membuktikan algoritma mana yang paling tangguh dalam menangani variabilitas dataset film, serta paling efektif dalam memberikan solusi personalisasi bagi pengguna.

Pendekatan pertama dan yang paling sederhana adalah *Demographic Filtering*. Metode ini sering digunakan sebagai model dasar (baseline) dalam banyak sistem. *Demographic Filtering* bekerja dengan mengelompokkan pengguna berdasarkan atribut umum atau merekomendasikan item yang populer secara global [5]. Implementasi teknis yang standar pada metode ini menggunakan formula *Weighted Rating* (seperti yang digunakan oleh IMDb), yang memberikan peringkat film berdasarkan akumulasi popularitas dan rata-rata penilaian. Keunggulan utamanya adalah kemampuannya mengatasi masalah pengguna baru (*cold-start*) dengan memberikan rekomendasi yang "aman" dan teruji kualitasnya, namun kelemahan fatalnya adalah sifatnya yang statis dan tidak mampu memberikan rekomendasi yang dipersonalisasi sesuai selera unik individu.

Berbeda dengan pendekatan demografis yang bersifat general, *Content-Based Filtering* menawarkan tingkat personalisasi yang lebih tinggi dengan menganalisis atribut intrinsik dari item itu sendiri. Metode ini dengan memanfaatkan teknik *Natural Language Processing* (NLP) seperti *Term Frequency-Inverse Document Frequency* (TF-IDF) untuk membedah deskripsi sinopsis, genre, dan aktor film. Premis dasarnya adalah jika seorang pengguna

menyukai film A, maka sistem akan merekomendasikan film B yang memiliki kemiripan atribut dengan film A. Metode ini sangat efektif untuk menemukan film dengan tema spesifik yang mungkin tidak populer namun sangat relevan dengan riwayat tontonan pengguna.

Namun, metode berbasis konten juga tidak lepas dari kelemahan. Hanafi et al. dan Li et al [6] [4]. mencatat bahwa metode ini rentan terhadap masalah overspecialization. Pengguna cenderung terjebak dalam "gelembung filter" (filter bubble), di mana sistem hanya merekomendasikan variasi dari apa yang sudah pernah mereka lihat. Misalnya, penggemar film aksi hanya akan terus disuguhi film aksi, menghilangkan unsur serendipity atau kejutan penemuan hal baru yang berbeda dari kebiasaan mereka namun tetap menarik. Hal ini membatasi eksplorasi pengguna terhadap katalog film yang lebih luas.

Untuk mengatasi keterbatasan tersebut, pendekatan Collaborative Filtering (CF) hadir sebagai solusi yang lebih dinamis dengan memanfaatkan konsep "kebijaksanaan orang banyak" (*wisdom of the crowd*). Kurniawan et al. [7] menjelaskan bahwa CF memprediksi preferensi pengguna berdasarkan pola interaksi komunitas[5]. Jika Pengguna A dan Pengguna B memiliki pola rating yang mirip di masa lalu, sistem berasumsi bahwa mereka akan memiliki selera yang sama di masa depan. Keunggulan utama CF adalah kemampuannya memberikan rekomendasi lintas-genre yang tidak terduga karena didasarkan pada perilaku manusia yang kompleks, bukan sekadar kemiripan teks.

Dalam implementasinya, algoritma CF menghadapi tantangan skalabilitas dan data yang jarang (*sparsity*). Untuk itu, teknik Matrix Factorization seperti Singular Value Decomposition (SVD) menjadi standar emas dalam industri. Fulzele et al. membuktikan bahwa SVD mampu mendekomposisi matriks interaksi pengguna-item menjadi faktor-faktor laten yang merepresentasikan fitur tersembunyi dari preferensi pengguna [8]. Penelitian Muliadi dan Lestari [5] juga mendukung bahwa pendekatan berbasis model seperti SVD menghasilkan akurasi prediksi yang lebih stabil dibandingkan pendekatan berbasis tetangga

(neighborhood-based) tradisional, terutama pada dataset dengan densitas rating yang rendah.

Meskipun berbagai metode telah tersedia, terdapat kesenjangan penelitian (research gap) dalam literatur saat ini. Sebagian besar penelitian terkini, seperti yang dilakukan Sharma et al. [9], cenderung langsung berfokus pada pengembangan sistem Hybrid yang kompleks tanpa terlebih dahulu melakukan komparasi mendalam mengenai kinerja fundamental dari ketiga lapisan algoritma penyusunnya secara terpisah pada dataset yang sama. Padahal, pemahaman mengenai trade-off antara akurasi (RMSE) dan karakteristik rekomendasi dari tiap metode sangat penting sebelum merancang sistem yang lebih kompleks.

Berdasarkan uraian di atas, penelitian ini bertujuan untuk merancang dan membandingkan secara komparatif kinerja dari tiga algoritma sistem rekomendasi: *Demographic Filtering* (berbasis *Weighted Rating*), *Content-Based Filtering* (berbasis TF-IDF dan Metadata), dan *Collaborative Filtering* (berbasis SVD).

1.2 Rumusan Masalah

1. Bagaimana merancang dan membangun sistem rekomendasi film yang menerapkan metode *Demographic Filtering*, *Content-Based Filtering*, dan *Collaborative Filtering* menggunakan dataset TMDb?
2. Bagaimana kinerja metode *Collaborative Filtering* menggunakan algoritma SVD (*Singular Value Decomposition*) dalam memprediksi rating pengguna yang diukur menggunakan metrik RMSE (*Root Mean Square Error*) dan MAE (*Mean Absolute Error*)?
3. Bagaimana karakteristik dan perbedaan hasil rekomendasi yang diberikan oleh pendekatan *Demographic Filtering* (sebagai *baseline non-personal*), *Content-Based Filtering* (berbasis kemiripan atribut film), dan *Collaborative Filtering* (berbasis pola interaksi pengguna) dalam berbagai skenario profil pengguna?

1.3 Rumusan Masalah

Lingkup penelitian dibatasi pada komparasi kinerja algoritma *Demographic Filtering* (Weighted Rating), *Content-Based* (TF-IDF/Cosine Similarity), dan *Collaborative Filtering* (SVD Scipy) menggunakan dataset TMDb 5000 dan data rating. Evaluasi berfokus pada akurasi prediksi menggunakan metrik RMSE dan MAE, dengan titik berat pada logika pemrosesan data (*backend*) tanpa membahas aspek antarmuka pengguna maupun keamanan jaringan.

1.4 Tujuan Penelitian

1. Mengimplementasikan metode *Demographic*, *Content-Based*, dan *Collaborative Filtering* dalam sebuah sistem rekomendasi film yang fungsional.
2. Mengukur tingkat akurasi prediksi metode *Collaborative Filtering* berbasis SVD untuk mengetahui tingkat error sistem dalam memprediksi preferensi pengguna.
3. Menganalisis perbandingan efektivitas rekomendasi, di mana *Content-Based* efektif untuk menangani item baru (*cold-start*), sedangkan *Collaborative Filtering* efektif untuk memberikan rekomendasi yang dipersonalisasi berdasarkan riwayat rating.
4. Implementasi *Collaborative Filtering* pada sistem rekomendasi film dalam bentuk aplikasi

1.5 Manfaat Penelitian

A. Manfaat Teoritis

1. Memberikan bukti empiris mengenai kinerja algoritma SVD (*Singular Value Decomposition*) dalam menangani dataset rating film skala kecil hingga menengah.
2. Menjadi referensi akademis bagi penelitian selanjutnya yang ingin mengembangkan sistem rekomendasi hibrida atau mengkomparasi metode *filtering* pada domain hiburan.

B. Manfaat Praktis

1. Menyediakan solusi algoritma yang dapat membantu pengguna mengatasi masalah *information overload* dalam memilih film yang sesuai dengan preferensi mereka.
2. Memberikan wawasan teknis bagi pengembang aplikasi *streaming* mengenai penerapan *preprocessing data* teks (seperti *Metadata Soup*) untuk meningkatkan relevansi rekomendasi berbasis konten.

1.6 Sistematika Penelitian

Sistematika penulisan dalam penyusunan laporan skripsi terdiri dari 5 bab, masing-masing diuraikan sebagai berikut:

Pada bab ini berisi mengenai latar belakang masalah, rumusan masalah, batasan masalah, maksud dan tujuan penelitian, metode penelitian dan sistematika penulisan yang digunakan untuk skripsi ini.

BAB II LANDASAN TEORI

Pada bab landasan teori merupakan tinjauan pustaka berupa tem yang pernah diteliti sebelumnya. Uraian teori teori yang mendasari pembahasan terperinci yang berhubungan dengan penelitian.

BAB III METODE PENELITIAN

Dalam bab ini menjelaskan tentang langkah-langkah yang dilakukan oleh penulis untuk melaksanakan dan menyelesaikan penelitian ini.

BAB IV IMPLEMENTASI DAN PEMBAHASAN

Bab ini berisi mengenai bagaimana menyelesaikan masalah yang telah dirumuskan berdasarkan langkah-langkah yang disebutkan pada Bab III, analisis, perancangan, dan pembahasan mengenai hasil penelitian.

BAB V PENUTUP

Bab ini membahas kesimpulan dan saran yang dapat diambil oleh peneliti yang telah dilakukan dan telah dibahas pada bab-bab sebelumnya.

REFERENSI

Bagian ini berisi daftar referensi yang telah digunakan dalam penulis laporan skripsi.

