

BAB I

PENDAHULUAN

1.1 Latar Belakang

Rekomendasi menjadi topik penelitian yang semakin penting dan menarik perhatian para peneliti. Sistem ini bertujuan untuk secara otomatis menyarankan produk atau layanan kepada *user* berdasarkan preferensi dan perilaku mereka, sehingga mengurangi kebutuhan untuk menjelajahi kumpulan data yang besar secara manual [1]. Rekomendasi telah menjadi bagian integral dari kehidupan dan banyak digunakan dalam berbagai aplikasi [2]. Algoritma rekomendasi *collaborative filtering* adalah salah satu algoritma rekomendasi yang paling terkenal, dengan fungsi utama untuk melakukan prediksi dan memberikan rekomendasi [3]. *Collaborative filtering* merupakan salah satu algoritma dalam sistem rekomendasi yang sederhana dan efisien, namun masih memiliki masalah seperti *cold-start*, kurangnya akurasi rekomendasi, serta ketidakmampuannya dalam menangkap hubungan kompleks antara *user* dan produk [1]. Permasalahan *cold-start* dalam *collaborative filtering* terjadi ketika *user* tidak memiliki riwayat penilaian sebelumnya. Akibatnya, rekomendasi tidak dapat diberikan bagi *user* yang riwayat penilaiannya tidak tersedia. Selain itu, tantangan utama lainnya adalah permasalahan *first-rater*, di mana suatu item baru tidak dapat direkomendasikan hingga memperoleh penilaian awal dari *user* [4]. Oleh karena itu, masalah *cold-start* masih perlu diperbaiki [5].

Sebagai teknologi penting dalam rekomendasi, *collaborative filtering* saat ini dibagi menjadi dua metode *user-based* dan *item-based*. Metode *user-based* mengevaluasi preferensi *user* berdasarkan pencarian informasi, ulasan, pembelian, dan data produk. Sementara itu, metode *item-based* digunakan untuk merekomendasikan item serupa kepada *user* yang mungkin menyukai produk tertentu, di mana hubungan antar *user* diperoleh dengan menghitung perbedaan penilaian *user* terhadap berbagai item. Namun dari kedua metode masih memiliki batasan tertentu, tidak dapat memenuhi keragaman dan kebaruan rekomendasi [6]. Banyak penelitian telah dilakukan dalam bidang *collaborative filtering* untuk rekomendasi [7]. Oleh karena itu, salah satu tantangan dalam penelitian adalah penggunaan algoritma *machine learning* untuk meningkatkan akurasi rekomendasi [8]. Pendekatan berbasis *hybrid* dalam *Collaborative Filtering* telah digunakan untuk mempersonalisasi model [7].

Matrix factorization adalah teknik populer yang digunakan dalam *collaborative filtering* [2]. *Matrix factorization* memecah matriks menjadi dua matriks peringkat lebih rendah, yang dapat digunakan untuk memprediksi penilaian *user* terhadap item [2], [9]. *Matrix factorization* telah banyak digunakan dalam penyaringan kolaboratif, dan beberapa studi telah mengeksplorasi penerapan *matrix factorization* dalam rekomendasi. Salah satu pendekatan yang digunakan adalah rekomendasi berbasis *hybrid*, yang mengintegrasikan *matrix factorization* dengan algoritma lain [2]. Seperti *extreme gradient boosting* (XGBoost) yang telah digunakan dalam rekomendasi secara berkelanjutan untuk memberikan rekomendasi yang dipersonalisasi dan relevan bagi *user*. Teknik ini sangat sesuai untuk tugas *collaborative filtering* [5], dimana XGBoost berperan untuk memprediksi peringkat atau skor preferensi *user* terhadap suatu item. Algoritma ini mengintegrasikan teknik pembelajaran berbasis *multi-classifiers* dan XGBoost yang digunakan untuk membangun ratusan pohon keputusan guna mempercepat proses pelatihan serta meningkatkan akurasi rekomendasi [9]. Kinerja XGBoost yang menjanjikan ini mendorong penelitian lebih lanjut, seperti pengembangan arsitektur berbasis *hybrid* baru yang diusulkan dalam penelitian ini. Arsitektur *hybrid* ini secara empiris terbukti melampaui batasan parameter eksperimental yang ditetapkan, sehingga mengindikasikan potensinya sebagai kandidat yang kuat untuk diimplementasikan dalam pengembangan rekomendasi [10].

Model berbasis *embedding* telah menjadi teknologi terkini dalam *collaborative filtering* selama lebih dari satu dekade ini [12]. Di antara model tersebut, metode *Multi-Layer Perceptron* (MLP) telah banyak diterapkan dalam rekomendasi. MLP merupakan *feedforward neural network* yang terdiri dari banyak lapisan tersembunyi dan telah banyak diterapkan dalam rekomendasi karena efisiensinya. Untuk memecahkan masalah di atas, sejumlah eksperimen telah dilakukan pada kumpulan data dunia nyata. Penelitian ini mencakup studi tentang interaksi antara *user identifier* dan *book identifier* dalam layanan rekomendasi pendidikan. Tujuan utamanya adalah untuk menganalisis efisiensi model yang digunakan dan membandingkannya dengan metode yang ada. Hasilnya menunjukkan peningkatan efisiensi rekomendasi, serta mengurangi masalah *cold-start* dan kelangkaan data [13].

Meskipun pendekatan *hybrid* telah banyak diterapkan dalam rekomendasi, sebagian besar studi terdahulu hanya mengombinasikan dua metode, seperti *Matrix Factorization* dengan *Feedforward Neural Network* atau *Matrix Factorization* dengan XGBoost. Hingga saat ini, sangat sedikit penelitian yang secara terpadu mengintegrasikan ketiga komponen tersebut ke dalam satu kerangka model *hybrid*. Kesenjangan ini membuka peluang strategis

untuk mengeksplorasi efektivitas kombinasi *Matrix Factorization* sebagai pengekstraksi faktor laten, *Feedforward Neural Network* untuk memodelkan interaksi non-linear yang kompleks, serta *XGBoost* sebagai pemodel prediksi akhir yang memanfaatkan fitur-fitur buatan (*handcrafted features*) secara optimal. Penelitian ini difokuskan pada pengembangan model rekomendasi *hybrid* yang menggabungkan metode *Collaborative Filtering* dengan pendekatan *Matrix Factorization*, *Feedforward Neural Network* berbasis *Multi-Layer Perceptron* dan *XGBoost*. Pengujian dilakukan secara komprehensif terhadap dua *dataset* berbeda, yaitu film dan hotel, untuk menguji efektivitas model dalam mengatasi permasalahan *cold-start* dan *first-rater*, sekaligus meningkatkan akurasi rekomendasi dalam domain berbeda. Dengan demikian, penelitian ini diharapkan tidak hanya memberikan kontribusi signifikan dalam penyempurnaan rekomendasi yang lebih personal dan relevan, tetapi juga membuka peluang bagi pengembangan algoritma *machine learning* yang lebih adaptif dan inovatif di era digital.

Dari permasalahan dan kajian literatur yang telah dijabarkan, maka dilakukan penelitian dengan mengambil judul “**Optimisasi Collaborative Filtering Berbasis Hybrid menggunakan Matrix Factorization, Feedforward Neural Network dan XGBoost**”.

1.2 Rumusan Masalah

Berdasarkan uraian pada latar belakang, rumusan masalah dalam penelitian ini adalah sebagai berikut: Bagaimana cara mengatasi permasalahan *cold-start* dan *first-rater* pada rekomendasi berbasis *Collaborative Filtering*? Sejauh mana efektivitas model rekomendasi *hybrid* yang mengintegrasikan *Collaborative Filtering*, *Matrix Factorization*, *Feedforward Neural Network* berbasis *Multi-Layer Perceptron*, dan *XGBoost* dalam mengatasi permasalahan tersebut? Selain itu, bagaimana implementasi model tersebut dapat meningkatkan akurasi rekomendasi melalui pengujian pada dua *dataset* yang berbeda, yaitu film dan hotel?

1.3 Tujuan

Tujuan penelitian ini adalah untuk mengembangkan dan mengukur akurasi model rekomendasi berbasis *hybrid* yang mengintegrasikan metode *Collaborative Filtering* dengan algoritma *Matrix Factorization*, *Feedforward Neural Network* berbasis *Multi-Layer Perceptron* dan *XGBoost* dalam mengatasi *cold-start* dan *first-rater* guna meningkatkan akurasi rekomendasi berdasarkan hasil pengujian pada masing-masing *dataset*.

1.4 Manfaat

Manfaat dari penelitian ini adalah sebagai berikut:

1. Model yang dikembangkan dapat memberikan kontribusi signifikan terhadap pengembangan rekomendasi melalui optimisasi metode *Collaborative Filtering* berbasis model, khususnya dengan menggabungkan *Matrix Factorization*, *Feedforward Neural Network* berbasis *Multi-Layer Perceptron* dan XGBoost, dengan tujuan meningkatkan akurasi rekomendasi.
2. Menjadi referensi bagi penelitian selanjutnya yang ingin mengeksplorasi pendekatan berbasis *hybrid* dalam sistem rekomendasi guna meningkatkan akurasi rekomendasi.

1.5 Ruang Lingkup

Ruang lingkup penelitian ini mencakup beberapa aspek sebagai berikut:

1. Penelitian ini difokuskan pada pengembangan model rekomendasi berbasis *hybrid* yang mengintegrasikan pendekatan *Collaborative Filtering*, *Matrix Factorization*, *Feedforward Neural Network* berbasis *Multi-Layer Perceptron* dan XGBoost untuk mengatasi keterbatasan dari masing-masing metode.
2. Penelitian dilakukan menggunakan dua jenis *dataset* publik, yaitu dataset film dan hotel, yang mewakili kategori rekomendasi yang berbeda. Hal ini bertujuan untuk menguji sejauh mana model yang diusulkan dapat diterapkan di berbagai konteks.
3. Penelitian ini berfokus pada rekomendasi menggunakan analisis teknikal dan mengabaikan aspek-aspek fundamental pada *dataset*.
4. Penelitian ini tidak membahas aspek bisnis, sosial, atau psikologis dari rekomendasi, serta tidak mencakup optimasi performa *runtime* secara mendalam seperti distribusi sistem aplikasi.
5. Pengujian pada model menggunakan MAE, MSE, RMSE dan R² Score untuk mengukur kinerja dan akurasi model yang diajukan.