

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pada era digital saat ini, perkembangan teknologi informasi memberikan dampak signifikan terhadap bidang literasi dan penyediaan konten bacaan digital. Kemudahan akses terhadap novel melalui berbagai *platform* daring menyebabkan jumlah pilihan bacaan semakin banyak dan beragam. Novel tidak hanya berfungsi sebagai hiburan, tetapi juga sebagai media pembelajaran dan pengembangan imajinasi. Namun, melimpahnya pilihan novel justru menyulitkan pembaca dalam menentukan bacaan yang sesuai dengan minat dan preferensinya.

Pada *platform* digital, tersedia ribuan hingga jutaan judul novel dengan variasi *Genre*, penulis, alur cerita, dan gaya bahasa. Kondisi tersebut membuat proses pemilihan bacaan secara manual menjadi kurang efisien dan berpotensi menghasilkan pilihan yang tidak sesuai dengan kebutuhan pengguna. Oleh karena itu, diperlukan sistem yang mampu membantu pembaca dalam memperoleh rekomendasi novel secara lebih terarah dan relevan (Kheng dkk., 2025)

Penelitian terdahulu mengenai sistem rekomendasi bacaan umumnya masih berfokus pada buku secara umum dan menggunakan pendekatan tunggal, seperti *Content-Based Filtering* dengan metode *TF-IDF* atau teknik kesamaan konten lainnya (Ardiansyah dkk., 2023; Ridhwanullah dkk., 2024). Pendekatan berbasis konten tersebut dinilai kurang optimal dalam menangani teks naratif panjang seperti sinopsis novel, sementara pendekatan berbasis kolaboratif sering menghadapi permasalahan cold start ketika data pengguna masih terbatas.

Penelitian internasional oleh (Li dkk., 2022) menegaskan bahwa sistem rekomendasi novel perlu mempertimbangkan karakteristik khusus novel dan perilaku pembaca untuk menghasilkan rekomendasi yang relevan. Namun, penelitian tersebut juga menunjukkan bahwa penggabungan informasi konten novel dengan preferensi pembaca masih belum banyak diterapkan secara bersamaan. Sebagian besar penelitian masih menggunakan teknik pengolahan teks konvensional sehingga kualitas rekomendasi belum optimal. Hal ini menunjukkan

bahwa kajian sistem rekomendasi yang secara khusus berfokus pada novel masih relatif terbatas.

Pendekatan berbasis konten umumnya hanya memanfaatkan kemiripan teks dan *Genre* sehingga rekomendasi yang dihasilkan cenderung kurang bervariasi, sedangkan *collaborative filtering* sering mengalami permasalahan cold start dan *Sparsity* ketika data interaksi pengguna masih terbatas (Ardiansyah dkk., 2023; Nesmaoui dkk., 2023). Oleh karena itu, pendekatan *Hybrid* banyak digunakan karena mampu menggabungkan keunggulan kedua metode tersebut untuk meningkatkan kualitas rekomendasi. Integrasi metode *BM25* dan *Matrix Factorization* dinilai mampu menghasilkan rekomendasi yang lebih relevan dan personal sesuai preferensi pengguna (Maramis dkk., 2025; Xu dkk., 2025).

Berdasarkan kondisi tersebut, penelitian ini berfokus pada novel sebagai objek kajian, bukan buku secara umum, dengan memperhatikan karakteristik teks naratif dan preferensi pembaca. Penelitian ini menerapkan pendekatan *Hybrid* dengan mengombinasikan metode *Best Matching 25 (BM25)* untuk mengukur relevansi konten teks novel dan *Matrix Factorization* untuk memodelkan preferensi laten pembaca berdasarkan data *Rating*. Pendekatan ini diharapkan mampu mengatasi keterbatasan metode tunggal yang banyak digunakan pada penelitian sebelumnya. Berdasarkan permasalahan tersebut, penelitian ini mengusulkan sistem rekomendasi novel menggunakan pendekatan *Hybrid* yang mengombinasikan metode *Best Matching 25 (BM25)* dan *Matrix Factorization*. Kombinasi kedua metode tersebut diharapkan mampu menghasilkan rekomendasi novel yang lebih relevan, personal, dan sesuai dengan preferensi pembaca dibandingkan pendekatan tunggal.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka rumusan masalah dalam penelitian ini adalah sebagai berikut.

1. Bagaimana hasil implementasi pendekatan *Hybrid Recommender System* yang mengombinasikan metode *Best Matching 25 (BM25)* dan *Matrix Factorization* dalam sistem rekomendasi novel?

2. Bagaimana kualitas hasil rekomendasi novel yang dihasilkan dari penerapan pendekatan *Hybrid* tersebut berdasarkan kesesuaian dengan minat dan preferensi pembaca?

1.3 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah sebagai berikut.

1. Mengimplementasikan pendekatan *Hybrid Recommender System* yang mengombinasikan metode *Best Matching 25 (BM25)* dan *Matrix Factorization* dalam sistem rekomendasi novel.
2. Mengevaluasi kualitas hasil rekomendasi novel yang dihasilkan dari penerapan pendekatan *Hybrid* berdasarkan kesesuaian dengan minat dan preferensi pembaca.

1.4 Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini adalah sebagai berikut.

1. Menambah wawasan dan referensi keilmuan dalam bidang sistem rekomendasi, khususnya penerapan pendekatan *Hybrid Recommender System* menggunakan metode *BM25* dan *Matrix Factorization*.
2. Menjadi acuan bagi penelitian selanjutnya yang mengembangkan sistem rekomendasi berbasis konten teks dan data *Rating* pengguna.
3. Memberikan gambaran penerapan sistem rekomendasi novel yang mampu menghasilkan rekomendasi yang lebih relevan dan sesuai dengan preferensi pembaca, sehingga dapat membantu pengguna dalam menentukan bacaan yang tepat.

1.5 Batasan Penelitian

Penelitian ini difokuskan pada beberapa batasan ruang lingkup sebagai berikut.

1. Penelitian ini menggunakan dataset buku umum dari *platform* Kaggle dan bersifat *open source* yang di dalamnya mencakup berbagai jenis bacaan. Namun, pada penelitian ini hanya item buku yang dikategorikan sebagai novel yang digunakan dalam proses analisis dan rekomendasi. Penentuan kategori novel dilakukan berdasarkan atribut *Genre* dan deskripsi sinopsis yang tersedia pada

dataset, sehingga item non-novel tidak disertakan dalam proses pemodelan sistem rekomendasi.

2. Dataset yang digunakan diambil dari dua sumber yang berbeda, yaitu dataset *Goodbooks-10k* yang terdiri dari sekitar 10.000 data buku dan 981.756 data *Rating*, serta dataset *Goodreads-books* yang berisi informasi konten buku berupa judul, penulis, *Genre*, dan sinopsis dengan jumlah data lebih dari 20.000 buku. Kedua dataset tidak digabungkan secara keseluruhan, tetapi disesuaikan melalui pencocokan *book_id* untuk memastikan data *Rating* yang digunakan memiliki pasangan pada dataset konten novel.
3. Sistem rekomendasi dalam penelitian ini dibangun menggunakan dua metode, yaitu *Best Matching 25 (BM25)* untuk rekomendasi berbasis konten dan *Matrix Factorization* untuk rekomendasi berbasis kolaboratif.
4. Sistem rekomendasi yang dibangun menghasilkan daftar rekomendasi novel dan diimplementasikan dalam bentuk *Website* berbasis *Flask* sebagai media penyajian hasil rekomendasi kepada pengguna. Sistem tidak terhubung dengan layanan eksternal, situs pihak ketiga, maupun *Application Programming Interface (API)* lainnya.