

BAB I PENDAHULUAN

1.1 Latar Belakang

Kemajuan teknologi perpustakaan digital telah mengubah cara pengguna mengakses dan mencari koleksi buku secara daring. Laporan *OverDrive* mencatat bahwa peminjaman konten digital mencapai 555 juta pada tahun 2022, meningkat 10% dari tahun sebelumnya [1]. Pada tahun 2023, jumlah tersebut kembali naik menjadi 662 juta atau bertambah 19%, yang menunjukkan peningkatan minat dan pemanfaatan terhadap layanan buku digital [2].

Namun, bertambahnya koleksi digital juga menimbulkan tantangan baru bagi pengguna. Banyak dari mereka kesulitan menemukan buku yang benar-benar sesuai kebutuhan karena sistem rekomendasi yang tersedia masih menghasilkan saran yang umum dan kurang relevan [3]. Kondisi ini terjadi karena informasi konten buku belum dimanfaatkan secara maksimal untuk menentukan kesesuaian rekomendasi.

Content-Based Filtering (CBF) menjadi pendekatan yang tepat untuk membantu mengatasi permasalahan tersebut. Metode ini bekerja dengan menganalisis kemiripan karakteristik item menggunakan representasi teks, seperti pembobotan *Term Frequency-Inverse Document Frequency (TF-IDF)* dan perhitungan jarak *Cosine Similarity* [4]. Pendekatan ini terbukti efektif dalam memberikan rekomendasi yang relevan berdasarkan deskripsi item dan preferensi pengguna tanpa bergantung pada data interaksi pengguna lain, sehingga mampu menangani masalah pada item baru yang belum memiliki *rating* [5].

Namun, pengukuran kesamaan saja seringkali memiliki keterbatasan, seperti rekomendasi yang terlalu monoton (*overspecialization*) atau kurang akurat dalam memprediksi preferensi yang kompleks [5]. Diperlukan model yang mampu mengevaluasi relevansi secara lebih mendalam. Penelitian terbaru menunjukkan bahwa algoritma *Extreme Gradient Boosting (XGBoost)* mampu meningkatkan presisi rekomendasi secara signifikan dibandingkan metode konvensional dengan

cara mempelajari pola fitur yang lebih rumit [6]. Oleh karena itu, kombinasi antara ekstraksi fitur *TF-IDF*, pengukuran *Cosine Similarity*, dan klasifikasi relevansi menggunakan *XGBoost* diharapkan dapat meningkatkan akurasi rekomendasi serta memberikan pengalaman pencarian buku yang lebih adaptif bagi pengguna perpustakaan digital.

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah yang telah diuraikan sebelumnya, maka inti permasalahan yang akan dikaji dalam penelitian ini dapat dirumuskan ke dalam beberapa pertanyaan penelitian sebagai berikut:

1. Bagaimana karakteristik metadata buku (judul, penulis, genre, dan deskripsi) mempengaruhi kualitas kemiripan rekomendasi pada metode *Content-Based Filtering*?
2. Bagaimana performa algoritma *XGBoost* dalam mengklasifikasikan relevansi rekomendasi buku?
3. Seberapa efektif integrasi *Content-Based Filtering* dan *XGBoost* dalam meningkatkan ketepatan rekomendasi buku dibandingkan pemetaan kemiripan konten tanpa model klasifikasi?

1.3 Batasan Masalah

Agar pembahasan penelitian ini lebih terarah, maka penulis menetapkan batasan-batasan masalah sebagai berikut:

1. Data yang digunakan terbatas pada atribut konten buku seperti judul, penulis, genre, dan deskripsi tanpa melibatkan data perilaku pengguna seperti *rating* atau *histori bacaan*.
2. Sistem rekomendasi yang dibangun tidak mempertimbangkan faktor seperti tren bacaan terkini, buku yang sedang populer, atau perubahan preferensi pengguna seiring waktu, model hanya menggunakan data statis yang tersedia dalam dataset saat penelitian dilakukan.
3. Evaluasi penelitian difokuskan pada matrik klasifikasi (*accuracy*, *precision*, *recall*, *f1-score*) dan tidak mencakup pengujian *usability* atau evaluasi pengalaman pengguna secara langsung.

1.4 Tujuan Penelitian

Mengacu pada rumusan masalah yang telah disusun, maka tujuan utama yang ingin dicapai melalui pelaksanaan penelitian ini adalah:

1. Mengimplementasikan *Content-Based Filtering* untuk menghasilkan rekomendasi buku berdasarkan kemiripan konten pada dataset perpustakaan digital.
2. Menganalisis performa algoritma *XGBoost* dalam mengklasifikasikan relevansi buku berdasarkan fitur teks yang dihasilkan *TF-IDF* dan nilai kemiripan konten dari *Cosine Similarity*.
3. Mengevaluasi efektivitas integrasi *Content-Based Filtering* dan *XGBoost* dalam meningkatkan ketepatan rekomendasi terhadap kebutuhan pengguna.

1.5 Manfaat Penelitian

Hasil dari penelitian ini diharapkan dapat memberikan kontribusi positif, baik secara teoritis maupun praktis bagi pengembangan ilmu pengetahuan dan teknologi, dengan rincian manfaat sebagai berikut:

1. Memberikan pemahaman mengenai faktor-faktor metadata buku yang paling berpengaruh dalam penentuan rekomendasi berbasis konten.
2. Memberikan gambaran mengenai tingkat performa metode *Content-Based Filtering* dengan algoritma *XGBoost* dalam menghasilkan rekomendasi buku yang relevan berdasarkan karakteristik konten buku.
3. Menyediakan dasar analisis yang dapat digunakan untuk pengembangan sistem rekomendasi berbasis konten di penelitian selanjutnya.

1.6 Sistematika Penulisan

Untuk memberikan gambaran yang jelas, terstruktur, dan menyeluruh mengenai alur pembahasan, maka materi pembahasan disusun ke dalam lima bab utama dengan rincian sebagai berikut:

1. BAB I Pendahuluan

Berisi latar belakang masalah, rumusan masalah, batasan masalah, tujuan

penelitian, manfaat penelitian, dan sistematika penulisan.

2. BAB II Tinjauan Pustaka

Berisi tinjauan teori terkait sistem rekomendasi, *Content-Based Filtering*, *machine learning*, *XGBoost*, serta penelitian terdahulu yang relevan. Bab ini juga mencakup tabel keaslian penelitian.

3. BAB III Metodologi Penelitian

Berisi gambaran umum objek penelitian, analisis masalah, alur metode, teknik ekstraksi fitur, proses pemodelan, serta rancangan sistem rekomendasi.

4. BAB IV Hasil dan Pembahasan

Berisi implementasi metode, hasil pengujian sistem rekomendasi, analisis performa model, pembahasan hasil, dan interpretasi metrik evaluasi.

5. BAB V Kesimpulan dan Saran

Berisi kesimpulan dari proses penelitian serta saran untuk pengembangan atau penelitian selanjutnya.