

BAB V PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan mengenai penerapan algoritma XGBoost pada metode Content-Based Filtering untuk klasifikasi relevansi buku, maka diperoleh kesimpulan sebagai berikut:

1. Karakteristik metadata buku seperti title, author, genre, dan description berpengaruh terhadap kualitas rekomendasi pada metode Content-Based Filtering. Atribut description menjadi sumber informasi utama karena memiliki representasi konten paling tinggi, sedangkan kombinasi seluruh atribut dalam fitur combined_text mampu meningkatkan representasi kesamaan antar buku.
2. Algoritma XGBoost mampu digunakan untuk mengklasifikasikan relevansi buku dengan performa yang cukup baik. Berdasarkan hasil evaluasi, model memperoleh rata-rata accuracy sebesar 0.87. Nilai precision pada kelas relevan sebesar 0.79 menunjukkan bahwa rekomendasi yang diberikan cukup tepat, meskipun nilai recall sebesar 0.49 menunjukkan masih terdapat beberapa buku relevan yang belum berhasil terdeteksi oleh model.
3. Integrasi metode Content-Based Filtering dan algoritma XGBoost efektif dalam meningkatkan proses analisis relevansi buku. Penggunaan fitur TF-IDF, cosine similarity, dan klasifikasi XGBoost memungkinkan sistem tidak hanya mengandalkan tingkat kemiripan konten, tetapi juga mempelajari pola relevansi yang lebih kompleks sehingga rekomendasi yang dihasilkan menjadi lebih akurat dan adaptif terhadap query pengguna.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa saran yang dapat dijadikan bahan pengembangan untuk penelitian selanjutnya, yaitu sebagai berikut:

1. Penelitian ini masih menggunakan data teks berbahasa Inggris pada proses preprocessing dan ekstraksi fitur. Oleh karena itu, penelitian selanjutnya dapat mengembangkan sistem agar dapat mendukung data dengan berbagai bahasa.

2. Pada penelitian ini, parameter model XGBoost masih menggunakan sebagian besar nilai bawaan (default). Oleh karena itu, penelitian selanjutnya dapat melakukan optimasi hyperparameter untuk meningkatkan performa model, khususnya pada nilai recall kelas relevan.
3. Representasi teks pada penelitian ini masih menggunakan metode TF-IDF yang berbasis frekuensi kata. Penelitian berikutnya dapat menggunakan metode representasi teks yang lebih modern seperti Word2Vec, atau BERT agar sistem mampu memahami konteks teks secara lebih mendalam.

