

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan pesat dalam sektor *e-commerce* telah memberikan kemudahan bagi konsumen dalam mengakses berbagai produk secara daring (Zuhdiansyah & Luthfiarta, 2024). Dengan banyaknya pilihan yang tersedia, sistem *e-commerce* berperan dalam meningkatkan efisiensi transaksi dan kenyamanan berbelanja. Namun, jumlah produk yang sangat besar juga dapat menyulitkan pengguna dalam menemukan item yang sesuai dengan preferensi mereka. Oleh karena itu, diperlukan suatu mekanisme yang dapat membantu proses pemilihan produk secara lebih efektif (Hartmann & Weißenberger, 2024). Salah satu solusi yang banyak diterapkan adalah dengan mengembangkan model rekomendasi.

Model rekomendasi merupakan sistem yang dirancang untuk mengidentifikasi preferensi pengguna dan memberikan saran produk yang sesuai dengan kebutuhan serta minat mereka (Suhailah & Hartatik, 2023). Beberapa pendekatan umum yang digunakan dalam pengembangan model rekomendasi meliputi *Collaborative Filtering* (CF) dan *Content-Based Filtering* (CBF). *Content-Based Filtering* bekerja dengan menganalisis karakteristik produk yang pernah disukai oleh pengguna, lalu merekomendasikan produk lain yang memiliki fitur serupa. Sistem ini mengandalkan informasi deskriptif dari produk, seperti kategori, merek, atau kata kunci dalam deskripsi. Meskipun efektif dalam menyarankan produk serupa, CBF memiliki keterbatasan karena hanya fokus pada

perilaku pengguna itu sendiri tanpa mempertimbangkan kesamaan antar pengguna, sehingga cenderung menghasilkan rekomendasi yang kurang bervariasi.

Collaborative Filtering menjadi salah satu metode yang paling banyak digunakan karena kemampuannya dalam menangkap pola konsumsi pengguna dengan menganalisis interaksi serta *rating* antar pengguna dan produk (Hasan, 2024). Proses kerja CF didasarkan pada asumsi bahwa pengguna yang memiliki pola interaksi serupa di masa lalu kemungkinan besar akan menyukai item yang sama di masa mendatang. Penelitian ini memilih CF sebagai pendekatan utama karena metode ini lebih fleksibel dalam menangani volume data besar pada lingkungan *e-commerce* dan memungkinkan integrasi dengan informasi tambahan, seperti ulasan pelanggan. Selain itu, CF dinilai lebih potensial dalam memberikan rekomendasi yang relevan dengan memanfaatkan data historis yang tersedia, terutama ketika deskripsi produk terbatas atau tidak seragam seperti pada sistem CBF.

Metode *Collaborative Filtering* menyimpan beberapa kelemahan. Permasalahan seperti *sparsity* atau minimnya data *rating* menghambat akurasi rekomendasi, ditambah adanya *gray sheep* pengguna dengan preferensi unik yang sulit dikelompokkan (Dervishaj & Cremonesi, 2022). Selain itu, metode ini sering kali hanya mengandalkan *rating* keseluruhan tanpa mempertimbangkan opini mendalam dalam ulasan pelanggan, sehingga rekomendasi kurang merefleksikan pengalaman pengguna secara komprehensif (Patil dkk., 2024) (Renukadevi dkk., 2024).

Upaya mengatasi keterbatasan tersebut dilakukan dengan menggabungkan *sentiment analysis* ke dalam CF. Penelitian (Karabila et al., 2023) mengintegrasikan CF dengan *BiLSTM* untuk mengekstraksi skor sentimen dari ulasan, dan terbukti meningkatkan akurasi. Keterbatasan *BiLSTM* terletak pada kemampuannya yang terbatas dalam memahami konteks kalimat karena sifat sekuensial dan arah pemrosesan yang terbatas (Rahman et al., 2024).

Penelitian ini mengusulkan integrasi model pra-latih *RoBERTa* sebuah model berbasis arsitektur *transformer* yang telah terbukti unggul dalam berbagai tugas pemrosesan bahasa alami dengan *Collaborative Filtering*. *RoBERTa* mampu menangkap konteks dan makna dalam teks ulasan dengan lebih akurat dibandingkan *BiLSTM* (X. Guo, 2024). Skor sentimen yang diekstraksi dari ulasan pelanggan menggunakan *RoBERTa* digunakan sebagai bobot tambahan dalam perhitungan kemiripan pengguna (Jahin et al., 2024). Dengan demikian, model rekomendasi tidak hanya mempertimbangkan *rating* numerik, tetapi juga opini dan emosi yang terkandung dalam ulasan, yang diharapkan dapat meningkatkan relevansi rekomendasi yang diberikan.

Evaluasi performa dilakukan dengan menggunakan metrik seperti *F1 Score*, *Accuracy*, *Precision*, *Recall*, dan *Root Mean Squared Error (RMSE)*, guna membandingkan efektivitas model integrasi terhadap model *Collaborative Filtering* konvensional tanpa dukungan *sentiment analysis*.

1.2 Rumusan Masalah

1. Bagaimana membangun model rekomendasi yang mengintegrasikan *sentiment analysis* berbasis *RoBERTa* dengan metode *Collaborative Filtering* untuk rekomendasi produk?
2. Bagaimana performa *sentiment analysis* terhadap kualitas rekomendasi produk dalam *e-commerce*?

1.3 Tujuan Penelitian

1. Membangun model rekomendasi produk dengan mengintegrasikan *sentiment analysis* berbasis *RoBERTa* dan metode *Collaborative Filtering* untuk rekomendasi produk.
2. Mengevaluasi pengaruh *sentiment analysis* terhadap kualitas rekomendasi dalam *e-commerce*.

1.4 Manfaat Penelitian

1. Bagi akademisi, penelitian ini memberikan kontribusi dalam pengembangan sistem rekomendasi *hybrid* yang menggabungkan *RoBERTa*-based *sentiment analysis* dan *User-based Collaborative Filtering*, sebagai pendekatan baru dalam kajian sistem cerdas berbasis data teks dan perilaku pengguna.
2. Bagi pengguna *e-commerce*, penelitian ini dapat dimanfaatkan untuk meningkatkan akurasi dan relevansi rekomendasi produk, dengan mempertimbangkan tidak hanya interaksi pengguna, tetapi juga sentimen dari ulasan pelanggan secara otomatis.

3. Bagi pengembang sistem, hasil penelitian ini dapat menjadi referensi dalam membangun sistem rekomendasi yang lebih responsif dan kontekstual, sehingga mampu meningkatkan kepuasan dan loyalitas pengguna dalam *platform* digital.

1.5 Batasan Penelitian

1. Data yang digunakan adalah dataset review pengguna *Amazon* kategori *All Beauty* periode 2021 – 2023. Dataset berjumlah 39993 ini mencakup informasi dasar terkait *rating*, *title*, *text*, *images*, *asin*, *parent_asin*, *user_id*, *timestamp*, *helpful_vote*, dan *verified_purchase*.
2. Penelitian ini akan mengintegrasikan model *RoBERTa* yang mengekstraksi dan meng*sentiment analysis* dari ulasan pengguna, serta model *Collaborative Filtering* yang digunakan untuk membuat rekomendasi.
3. Evaluasi performa model untuk *sentiment analysis* akan menggunakan *F1 Score*, *Accuracy*, *Precision* dan *Recall*, sedangkan dalam model rekomendasi akan menggunakan *Root Mean Squared Error (RMSE)*.