

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan pesat platform e-commerce telah mendorong peningkatan jumlah dan variasi produk yang tersedia bagi konsumen. Kondisi ini memicu fenomena kelebihan informasi (*information overload*), yaitu keadaan ketika jumlah informasi yang tersedia melebihi kemampuan konsumen dalam memprosesnya. Fenomena ini mencakup beberapa aspek, antara lain jumlah produk yang sangat banyak, kompleksitas atribut produk, serta kelelahan kognitif yang dapat menimbulkan kebingungan bagi konsumen (Qiu *et al.*, 2025). *Information overload* terbukti berdampak negatif terhadap pengalaman belanja pengguna, seperti meningkatnya kesulitan dalam pengambilan keputusan, kecenderungan melakukan pembelian impulsif, munculnya ketidakpuasan setelah pembelian, serta meningkatnya keinginan untuk membatalkan atau mengembalikan produk (Lv and Liu, 2022).

Platform e-commerce berskala besar menyediakan jutaan produk dari berbagai penjual, termasuk pelaku usaha kecil dan menengah, dalam ribuan kategori produk. Kondisi ini berpotensi memperparah fenomena *information overload* yang dialami pengguna. Teofilus *et al.* (2020) menunjukkan bahwa *information overload* berpengaruh signifikan terhadap kecenderungan pengguna untuk berhenti menggunakan platform e-commerce. Dalam kondisi tersebut, sistem rekomendasi menjadi komponen penting yang berfungsi membantu pengguna

menyaring informasi dan menemukan produk yang relevan di tengah banyaknya pilihan yang tersedia (Abdul Hussien, Rahma and Abdul Wahab, 2021).

Berbagai pendekatan sistem rekomendasi telah dikembangkan dalam ekosistem e-commerce, seperti pendekatan *collaborative filtering* (Rahmayuna, N. Harini, 2025), *content-based filtering* (Halimatussa'diyah and Dini, 2025), serta sistem rekomendasi *hybrid* yang mengombinasikan beberapa metode (Dinda Maristha, Santoso and Sari Dewi, 2021; Keikhosrokiani and Man Fye, 2024). Di antara pendekatan tersebut, *content-based filtering* banyak digunakan karena memanfaatkan atribut produk seperti nama, kategori, dan deskripsi, serta tidak bergantung pada data interaksi pengguna lain. Hal ini membuat pendekatan tersebut relatif lebih stabil ketika data pengguna terbatas (Pazzani and Billsus, 2007; Isinkaye, Folajimi and Ojokoh, 2015; Nurcahya and Supriyanto, 2020; Abdul Hussien, Rahma and Abdul Wahab, 2021).

Namun demikian, sistem rekomendasi menghadapi tantangan berupa kecenderungan popularitas (*popularity bias*), yaitu kondisi ketika sistem lebih sering merekomendasikan produk dengan tingkat popularitas tinggi berdasarkan jumlah interaksi dan ulasan, sementara produk dengan popularitas rendah cenderung jarang direkomendasikan. Akibatnya, produk dengan tingkat keterpaparan rendah memiliki peluang yang lebih kecil untuk ditemukan oleh pengguna, meskipun sebenarnya relevan bagi segmen pengguna tertentu (Qin, 2021; Carnovalini, Rodà and Wiggins, 2025).

Dalam platform e-commerce berskala besar, kecenderungan ini berdampak pada rendahnya keterpaparan produk milik penjual baru atau pelaku usaha kecil. Produk-produk tersebut belum tentu memiliki kualitas yang lebih rendah dibandingkan produk populer, melainkan kurang mendapatkan perhatian dari sistem rekomendasi karena minimnya riwayat interaksi. Kondisi ini menyebabkan ketimpangan distribusi rekomendasi dan visibilitas produk dalam platform, yang pada dasarnya berkaitan dengan mekanisme pemeringkatan dalam sistem — yaitu bagaimana sistem menentukan urutan produk yang ditampilkan kepada pengguna (Abdollahpouri *et al.*, 2021; Qin, 2021; Carnovalini, Rodà and Wiggins, 2025).

Berbagai pendekatan telah dikembangkan untuk mengurangi kecenderungan popularitas tersebut, seperti *re-ranking*, *post-processing*, serta pendekatan berbasis *deep learning* (Abdollahpouri, Burke and Mobasher, 2019; Abdollahpouri *et al.*, 2021a; Gupta *et al.*, 2021; Machado, Brandão and Marques-Neto, 2024). Namun, sebagian besar metode tersebut diterapkan pada pendekatan *collaborative filtering* dan cenderung memiliki kompleksitas yang tinggi. Sementara itu, penerapan mekanisme pengendalian popularitas pada sistem rekomendasi berbasis konten, khususnya pada domain produk e-commerce, masih relatif terbatas (Qin, 2021).

Berdasarkan permasalahan tersebut, penelitian ini mengusulkan pendekatan *hybrid content-based filtering* dengan mekanisme *popularity debiasing*. Pendekatan ini dirancang untuk mengurangi kecenderungan popularitas sekaligus meningkatkan keterpaparan produk *long-tail*, tanpa mengorbankan relevansi rekomendasi yang dihasilkan. Efektivitas pendekatan ini dievaluasi menggunakan

dataset Amazon Reviews 2023 kategori *Toys and Games* yang menyediakan data interaksi pengguna dan metadata produk yang lengkap untuk mendukung evaluasi secara komprehensif.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, permasalahan utama dalam sistem rekomendasi produk e-commerce adalah kecenderungan sistem dalam memprioritaskan produk dengan tingkat popularitas tinggi, sehingga menyebabkan rendahnya keterpaparan produk *long-tail* meskipun produk tersebut berpotensi relevan bagi pengguna. Selain itu, sebagian besar pendekatan mitigasi *popularity bias* yang ada masih berfokus pada *collaborative filtering* dan belum banyak diintegrasikan secara eksplisit ke dalam sistem rekomendasi berbasis konten.

Berdasarkan permasalahan tersebut, rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana kinerja sistem rekomendasi berbasis konten murni (*Pure Content-Based Filtering*) dalam menghasilkan rekomendasi yang relevan, dan seberapa besar kecenderungan *popularity bias* yang muncul secara alami pada pendekatan tersebut?
2. Sejauh mana pendekatan *Hybrid Content-Based Filtering* dengan mekanisme *popularity debiasing* mampu mengurangi *popularity bias* dan meningkatkan keterpaparan produk *long-tail*, serta bagaimana *trade-off* yang terjadi terhadap kualitas relevansi rekomendasi dibandingkan dengan *baseline* Pure CBF dan *Popularity-Based*?

1.3 Tujuan Penelitian

Penelitian ini bertujuan untuk mengkaji dan mengevaluasi kinerja pendekatan *content-based filtering* yang dikombinasikan dengan mekanisme pembobotan popularitas dan kualitas dalam sistem rekomendasi produk e-commerce. Kinerja sistem rekomendasi diukur menggunakan metrik akurasi dan metrik di luar akurasi (*beyond-accuracy*), dengan relevansi rekomendasi didefinisikan sebagai kesesuaian atribut produk. Secara khusus, tujuan penelitian ini adalah sebagai berikut :

1. Menganalisis kinerja sistem rekomendasi *Content-Based Filtering* dalam menghasilkan rekomendasi yang relevan, serta mengkaji seberapa besar kecenderungan *popularity bias* yang muncul secara alami pada pendekatan tersebut.
2. Mengevaluasi efektivitas pendekatan *Hybrid Content-Based Filtering* dengan mekanisme *popularity debiasing* dalam mengurangi *popularity bias* dan meningkatkan eksposur produk *long-tail*, serta membandingkan *trade-off* kinerja yang dihasilkan terhadap *baseline* Pure CBF dan *Popularity-Based*.

1.4 Batasan Masalah

Agar penelitian ini terfokus dan terarah, batasan masalah yang ditetapkan adalah sebagai berikut:

1. Penelitian menggunakan dataset Amazon Reviews 2023 kategori *Toys and Games* yang dikembangkan oleh *McAuley-Lab*, University of California San Diego sebagai sumber data utama.
2. Pendekatan yang digunakan adalah *Hybrid Content-Based Filtering* dengan mekanisme *Inverse Popularity Weighting* tanpa melibatkan *Collaborative Filtering*.
3. Data yang digunakan mencakup data interaksi pengguna (*rating* dan *timestamp*) serta metadata produk berupa atribut konten (judul, kategori, fitur, dan nama toko), jumlah ulasan, dan nilai rata-rata *rating*.
4. Evaluasi sistem dilakukan secara *offline* menggunakan protokol *last-out* dengan metrik akurasi (*Precision@K*, *Recall@K*, *NDCG@K*) dan metrik *beyond-accuracy* (*LT Coverage@K*, *Average Popularity@K*, *Diversity@K*, *Catalog Coverage@K*) pada saat nilai $K = 5, 10$, dan 20 .
5. Penelitian ini tidak melibatkan evaluasi langsung oleh pengguna (*user study*) maupun analisis dampak bisnis secara nyata.

1.5 Manfaat Penelitian

Secara akademik, penelitian ini diharapkan dapat menambah referensi ilmiah di bidang sistem rekomendasi e-commerce, khususnya terkait penerapan mekanisme *popularity debiasing* pada pendekatan *content-based filtering*, serta penggunaan kombinasi metrik akurasi dan *beyond-accuracy* dalam menilai *trade-off* kinerja sistem rekomendasi secara komprehensif.

Secara praktis, penelitian ini diharapkan dapat menjadi referensi bagi pengelola platform e-commerce dalam merancang sistem rekomendasi yang tidak

hanya mempertimbangkan relevansi produk, tetapi juga mampu meningkatkan keterpaparan produk *long-tail* secara lebih merata tanpa mengorbankan kualitas rekomendasi yang dihasilkan.