

BAB I PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi digital dalam beberapa tahun terakhir telah memicu peningkatan volume data dan konten pada berbagai platform, seperti layanan *streaming*, *e-commerce*, dan media sosial. Melimpahnya konten tersebut menyulitkan pengguna dalam menemukan item yang relevan secara cepat. Dalam konteks ini, sistem rekomendasi berperan penting sebagai mekanisme untuk menyaring dan menyajikan konten sesuai preferensi pengguna [1].

Selain permasalahan relevansi, sistem rekomendasi tradisional juga kurang mampu merepresentasikan perubahan preferensi pengguna dari waktu ke waktu. Model yang bersifat statis cenderung mengabaikan urutan interaksi pengguna, padahal informasi temporal dalam riwayat interaksi berperan penting dalam memahami dinamika preferensi [2].

Di sisi lain, sistem rekomendasi juga menghadapi permasalahan bias, terutama *popularity bias* dan *exposure bias*. *Popularity bias* menyebabkan item dengan jumlah interaksi tinggi secara historis lebih sering direkomendasikan, sementara *exposure bias* muncul karena item yang sering ditampilkan memiliki peluang lebih besar untuk kembali dipilih. Akibatnya, terjadi efek penguatan berulang (*feedback loop*) yang membuat item populer semakin dominan dalam daftar rekomendasi [3]; [4].

Dalam praktiknya, fenomena ini dapat diamati pada layanan *streaming*, di mana film dengan jumlah tontonan tinggi cenderung terus muncul pada halaman utama, sementara film kategori long-tail jarang terekspose. Hal serupa juga terjadi pada platform *e-commerce*, di mana produk dengan label “best seller” lebih sering direkomendasikan dibandingkan produk baru. Dampak dari kondisi tersebut adalah menurunnya keberagaman rekomendasi dan terbentuknya pengalaman pengguna yang homogen [5].

Perkembangan *Large Language Model* (LLM) membuka peluang baru

dalam sistem rekomendasi melalui kemampuannya dalam membangun representasi semantik yang kaya dan kontekstual. Berbeda dengan pendekatan konvensional yang memanfaatkan LLM sebagai komponen *reranking* terpisah, penelitian ini memandang LLM sebagai bagian integral dalam proses pembentukan representasi laten pengguna dan item secara *end-to-end* [6].

Melalui pendekatan *End-to-End Embedding Fusion*, embedding sekuensial yang dihasilkan oleh model numerik (SASRec) digabungkan secara langsung dengan embedding semantik dari LLM untuk membentuk representasi terpadu [7]. Pendekatan ini memungkinkan sistem rekomendasi untuk memanfaatkan informasi perilaku historis pengguna dan konteks semantik item secara simultan dalam satu proses pembelajaran terpadu, tanpa memisahkan tahap pemodelan dan penyusunan peringkat [8].

Sementara itu, pendekatan debiasing seperti Inverse Propensity Scoring (IPS) terbukti efektif dalam mengurangi bias eksposur [9], namun integrasinya dalam kerangka *end-to-end* yang menggabungkan model sekuensial dan LLM masih relatif terbatas. Oleh karena itu, penelitian ini berfokus pada evaluasi integrasi SASRec, LLM berbasis *End-to-End Embedding Fusion*, dan IPS secara kuantitatif guna menghasilkan sistem rekomendasi yang lebih akurat, adaptif, dan adil.

1.2 Rumusan Masalah

1. Bagaimana kinerja model sekuensial SASRec dalam menangkap pola urutan interaksi pengguna untuk sistem rekomendasi??
2. Bagaimana pengaruh integrasi *Large Language Model (LLM)* melalui pendekatan *End-to-End Embedding Fusion* terhadap kualitas rekomendasi pada sistem berbasis SASRec?
3. Bagaimana penerapan *Inverse Propensity Scoring (IPS)* dalam sistem rekomendasi *end-to-end* berbasis integrasi SASRec dan LLM untuk mitigasi bias popularitas dan bias eksposur?

1.3 Batasan Masalah

Untuk menjaga fokus dan kedalaman penelitian, maka penelitian ini dibatasi pada beberapa aspek sebagai berikut:

1. Data yang digunakan berupa data interaksi pengguna-item *implicit feedback* yang bersifat historis dan telah tersedia, tanpa melibatkan pengumpulan data secara langsung atau *real-time* dari pengguna.
2. Penelitian difokuskan pada pengembangan dan evaluasi sistem rekomendasi berbasis *sequential recommendation*, dengan SASRec digunakan sebagai model utama untuk memodelkan dinamika preferensi pengguna berdasarkan urutan interaksi historis.
3. *Large Language Model* (LLM) dimanfaatkan sebagai komponen pembentuk representasi semantik item yang diintegrasikan secara langsung dengan representasi sekuensial pengguna melalui pendekatan *end-to-end embedding fusion*, tanpa memisahkan proses pemodelan ke dalam tahapan rekomendasi yang terpisah.
4. Mekanisme pengurangan bias hanya menggunakan pendekatan *Inverse Propensity Scoring* (IPS) untuk menangani bias popularitas dan bias eksposur, tanpa membahas atau membandingkannya dengan metode debiasing lain.
5. Evaluasi kinerja sistem rekomendasi dilakukan menggunakan metrik evaluasi kuantitatif, yang mencakup metrik berbasis peringkat dan metrik berbasis klasifikasi, tanpa melibatkan evaluasi berbasis survei atau umpan balik langsung dari pengguna.

1.4 Tujuan Penelitian

1. Menganalisis kinerja model SASRec dalam memodelkan dinamika preferensi pengguna berdasarkan urutan interaksi historis.
2. Mengevaluasi pengaruh integrasi *Large Language Model* (LLM) melalui pendekatan *End-to-End Embedding Fusion* terhadap peningkatan kualitas

rekomendasi pada sistem berbasis SASRec.

3. Mengimplementasikan dan menganalisis penerapan Inverse Propensity Scoring (IPS) sebagai mekanisme mitigasi bias popularitas dan bias eksposur dalam sistem rekomendasi end-to-end berbasis integrasi SASRec dan LLM.

1.5 Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini ditinjau dari dua aspek, yaitu manfaat teoretis dan manfaat praktis, sebagai berikut:

1. Manfaat Teoretis

Hasil penelitian ini diharapkan dapat menjadi referensi dalam pengembangan sistem rekomendasi yang mengintegrasikan pemodelan preferensi sekuensial, representasi semantik, dan mekanisme mitigasi bias dalam satu kerangka pembelajaran terpadu.

2. Manfaat Praktis

Secara praktis, penelitian ini diharapkan dapat menjadi acuan dalam pengembangan sistem rekomendasi berbasis data interaksi historis pengguna yang mampu meningkatkan relevansi rekomendasi, mengurangi bias popularitas dan bias eksposur, serta mendukung pengalaman pengguna yang lebih personal dan seimbang.

1.6 Sistematika Penulisan

Sistematika penulisan skripsi ini disusun secara terstruktur dalam lima bab yang saling berkaitan untuk memberikan pemahaman menyeluruh tentang penelitian yang dilakukan. Adapun uraian singkat isi dari masing-masing bab adalah sebagai berikut:

BAB I PENDAHULUAN

Bab ini menguraikan latar belakang penelitian yang membahas perkembangan teknologi digital, permasalahan sistem rekomendasi tradisional, serta peluang pemanfaatan Large Language Model dalam sistem rekomendasi. Bab ini juga

memuat rumusan masalah yang mencakup evaluasi kinerja SASRec, pengaruh integrasi LLM melalui *End-to-End Embedding Fusion*, dan penerapan *Inverse Propensity Scoring* untuk mitigasi bias. Selain itu, bab ini menjelaskan batasan masalah penelitian, tujuan penelitian yang ingin dicapai, manfaat teoretis dan praktis dari penelitian, serta sistematika penulisan skripsi.

BAB II TINJAUAN PUSTAKA

terdahulu mengenai SASRec, pemodelan sekuensial, bias dalam sistem rekomendasi, *Inverse Propensity Scoring*, dan pemanfaatan Large Language Model. Bab ini juga memuat dasar-dasar teori yang digunakan, meliputi konsep sistem rekomendasi, pendekatan *collaborative filtering* berbasis model, *sequential recommendation*, arsitektur SASRec, integrasi representasi dalam sistem *end-to-end*, pemanfaatan LLM untuk representasi semantik, permasalahan bias dalam sistem rekomendasi, mekanisme *Inverse Propensity Scoring*, serta metrik evaluasi yang digunakan untuk mengukur kinerja sistem rekomendasi.

BAB III METODE PENELITIAN

Bab ini menjelaskan alur penelitian secara sistematis, dimulai dari tahap persiapan data menggunakan dataset MovieLens-1M, pra-pemrosesan data yang mencakup rating filtering, user dan item filtering, re-indexing, sequential split, dan perhitungan bobot IPS. Bab ini juga menguraikan skenario pengujian yang terdiri dari tiga tahap (SASRec Only, SASRec dan LLM, dan SASRec, LLM, dan IPS), implementasi model *end-to-end* yang mengintegrasikan *sequential modeling*, *semantic embedding*, *embedding fusion*, *scoring & prediction*, dan *IPS-weighted loss*. Selain itu, bab ini membahas metode evaluasi kinerja menggunakan metrik *ranking-based*, *classification-based*, dan *fairness metrics*, serta tahap analisis dan pembahasan hasil. Pada bagian akhir, bab ini memaparkan alat dan bahan penelitian yang mencakup spesifikasi platform komputasi, software, framework, dan library yang digunakan.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil implementasi dan evaluasi sistem rekomendasi yang

dikembangkan. Bab ini dimulai dengan penyajian statistik dan karakteristik *dataset* MovieLens-1M, hasil dari setiap tahapan pra-pemrosesan data, detail implementasi model end-to-end dengan konfigurasi parameter yang digunakan. Selanjutnya, bab ini menampilkan hasil evaluasi kinerja untuk setiap skenario pengujian menggunakan metrik HR@K, NDCG@K, AUC, serta *fairness metrics*. Bab ini juga menyajikan perbandingan antar skenario dan analisis mendalam mengenai kontribusi SASRec dalam menangkap pola urutan interaksi, pengaruh integrasi LLM terhadap kualitas rekomendasi, dan efektivitas penerapan IPS dalam mitigasi bias popularitas dan eksposur.

BAB V PENUTUP

Bab ini berisi kesimpulan dari hasil penelitian yang menjawab rumusan masalah, meliputi efektivitas model SASRec dalam memodelkan dinamika preferensi pengguna, kontribusi integrasi LLM melalui *End-to-End Embedding Fusion* terhadap peningkatan kualitas rekomendasi, serta keberhasilan penerapan IPS dalam meningkatkan *fairness* dengan *trade-off* akurasi yang minimal. Bab ini juga menyajikan saran untuk pengembangan penelitian selanjutnya, yang mencakup eksplorasi arsitektur LLM yang lebih besar, pengembangan mekanisme fusi yang lebih canggih, penerapan metode debiasing tambahan, evaluasi pada *dataset* yang lebih beragam, integrasi mekanisme *explainability*, dan evaluasi melalui user study dalam kondisi real-world.