

# Analisis Komparatif Embedding Semantik Berbasis Large Language Model Pada Sistem Rekomendasi Buku Serendipitous di Perpustakaan Kampus

*Enhancing Serendipitous Book Recommendation Using LLM-Based Embeddings in Sequential Recommendation Models for University Library*

Rahayu Kartika Sari \*<sup>1</sup>, Eka Dyar Wahyuni<sup>2</sup>, Amalia Anjani Arifiyanti<sup>3</sup>  
<sup>1,2,3</sup>Sistem Informasi, Universitas Pembangunan Nasional "Veteran" Jawa Timur  
E-mail: <sup>1</sup>22082010146@student.upnjatim.ac.id, <sup>2</sup>penulis.2@email.ac.id,  
<sup>3</sup>penulis.3@email.ac.id

## Abstrak

Fenomena kelebihan informasi (*information overload*) di perpustakaan akademik sering kali menyulitkan pengguna dalam menemukan buku yang relevan, sehingga berdampak pada rendahnya minat baca. Di sisi lain, sistem rekomendasi konvensional rentan terhadap filter bubbles dan mengalami penurunan performa pada kondisi *cold-start*. Penelitian ini mengusulkan sistem rekomendasi sekuensial berbasis *Self-Attention Based Sequential Recommendation (SASRec)* yang diintegrasikan dengan lima model embedding semantik, yaitu *Word2Vec*, *BERT Multilingual*, *OpenAI text-embedding-3-small*, *Gemini-embedding-001*, dan *Qwen3-Embedding-0.6B*, untuk menghasilkan rekomendasi yang akurat sekaligus bersifat serendipity. Selain itu, algoritma *Serendipity-Oriented Greedy (SOG)* re-ranking diterapkan untuk menyeimbangkan relevansi dan serendipity rekomendasi. Dataset penelitian terdiri atas 14.502 data buku dan 5.445 riwayat interaksi pengguna setelah proses pembersihan data. Evaluasi dilakukan pada tiga skenario pengujian, yaitu *all test set*, *warm test set*, dan *cold test set*, dengan membandingkan seluruh varian model sebelum dan sesudah proses re-ranking. Hasil penelitian menunjukkan bahwa integrasi embedding berbasis Large Language Model (LLM) secara konsisten meningkatkan performa dibandingkan SASRec standar dan embedding tradisional. Model *Qwen3-Embedding-0.6B* memberikan performa terbaik dengan peningkatan *HitRate@10* hingga 282,9% dan *NDCG@10* hingga 387,8%, serta mempertahankan stabilitas semantik pada skenario *cold-start* dengan nilai *UnSerendipity@K* sebesar 0,613. Penerapan *SOG* re-ranking menunjukkan adanya trade-off antara akurasi dan keberagaman rekomendasi dimana pembobotan ringan menghasilkan keseimbangan optimal, sedangkan pembobotan yang terlalu agresif menurunkan relevansi secara signifikan. Kontribusi utama penelitian ini terletak pada integrasi embedding LLM modern ke dalam arsitektur rekomendasi sekuensial untuk meningkatkan akurasi dan ketahanan *cold-start*, sekaligus mengevaluasi pengaruh strategi serendipity re-ranking terhadap keseimbangan relevansi dan keberagaman rekomendasi. Dengan demikian, penelitian ini berhasil menghasilkan sistem rekomendasi perpustakaan yang lebih cerdas, adaptif, dan seimbang antara akurasi serta serendipity.

**Kata kunci:** *Item Embedding, Large Language Model, SASRec, Serendipity, Sistem Rekomendasi, Perpustakaan*

## Abstract

The phenomenon of information overload in academic libraries often makes it difficult for users to discover relevant books, which may reduce reading interest. Conventional recommender systems are also prone to filter bubbles and tend to perform poorly under cold-start conditions. This study proposes a sequential recommendation system based on the *Self-Attention Based Sequential Recommendation (SASRec)* model integrated with five semantic

*embedding models, namely Word2Vec, BERT Multilingual, OpenAI text-embedding-3-small, Gemini-embedding-001, and Qwen3-Embedding-0.6B, to generate accurate and serendipitous recommendations. In addition, the Serendipity-Oriented Greedy (SOG) re-ranking algorithm is implemented to balance recommendation relevance and serendipity. The data set consists of 14,502 book records and 5,445 user interaction histories after the data cleaning process. Evaluation was conducted under three testing scenarios, namely the all-test set, warm test set, and cold test set, by comparing all model variants before and after the re-ranking process. The results show that the integration of Large Language Model (LLM)-based embeddings consistently improves performance compared to the standard SASRec model and traditional embeddings. Qwen3-Embedding-0.6B achieved the best performance, improving HitRate@10 by up to 282.9% and NDCG@10 by up to 387.8%, while maintaining semantic robustness in cold-start scenarios with an UnSerendipity@K score of 0.613. The implementation of SOG re-ranking reveals a direct trade-off between recommendation accuracy and diversity. Lightweight weighting provides the optimal balance, whereas overly aggressive weighting significantly reduces relevance. The main contribution of this study lies in integrating modern LLM embeddings into a sequential recommendation architecture to improve accuracy and cold-start robustness, while also evaluating the impact of serendipity-oriented re-ranking strategies on balancing recommendation relevance and diversity. Overall, this study demonstrates that modern LLM integration can produce a smarter, more adaptive, and more balanced library recommendation system in terms of both accuracy and serendipity.*

**Keywords:** *Item Embedding, Large Language Model, SASRec, Serendipity, Recommender System, Library*

## 1. PENDAHULUAN

Perkembangan teknologi digital telah meningkatkan ketersediaan dan aksesibilitas informasi secara signifikan, termasuk dalam bidang literasi dan perpustakaan. Jumlah koleksi buku yang terus bertambah memberikan kemudahan bagi pengguna untuk memperoleh sumber pengetahuan, baik dalam bentuk fisik maupun digital [1], [2]. Namun, melimpahnya informasi tersebut juga menimbulkan fenomena *information overload*, yaitu kondisi ketika pengguna mengalami kesulitan dalam menemukan informasi yang benar-benar relevan di antara banyaknya pilihan yang tersedia [3].

Salah satu pendekatan yang banyak digunakan untuk mengatasi permasalahan tersebut adalah sistem rekomendasi (*recommender system*) [4], [5]. Sistem rekomendasi mampu membantu pengguna menemukan item yang relevan berdasarkan preferensi dan pola interaksi pengguna sebelumnya [6], [7], [8], [9]. Dalam konteks perpustakaan, sistem rekomendasi dapat berperan sebagai pustakawan virtual yang membantu pengguna menemukan buku sesuai kebutuhan akademik maupun minat baca mereka. Akan tetapi, pendekatan rekomendasi konvensional seperti *Collaborative Filtering* (CF) dan *Content-Based Filtering* (CBF) masih memiliki berbagai keterbatasan. *Collaborative Filtering* rentan terhadap masalah *cold-start*, *data sparsity*, dan *popularity bias*, sedangkan *Content-Based Filtering* cenderung menghasilkan rekomendasi yang terlalu homogen sehingga memicu fenomena *filter bubble* dan *overspecialization* [10], [11], [12]. Permasalahan ini sering kali menghasilkan rekomendasi repetitif dan membatasi eksplorasi pengguna terhadap konten yang baru dan beragam [13].

Untuk mengatasi keterbatasan tersebut, penelitian terbaru menyoroti pentingnya integrasi aspek *serendipity* ke dalam algoritma rekomendasi [14], [15]. Dengan merekomendasikan item yang relevan, baru, dan tidak terduga secara bersamaan, sistem rekomendasi *serendipitous* dapat mendorong eksplorasi pengguna sekaligus meningkatkan keterlibatan pengguna [16]. Namun, mencapai keseimbangan yang efektif antara relevansi dan *novelty* (keterbaruan) masih menjadi tantangan yang kompleks, karena membutuhkan pemodelan preferensi pengguna yang akurat serta pemahaman mendalam terhadap konten item.

Selain aspek *serendipity*, tantangan lain dalam sistem rekomendasi adalah memahami perubahan preferensi pengguna yang bersifat dinamis. Preferensi pengguna terhadap buku dapat berubah seiring waktu mengikuti konteks pembelajaran, kebutuhan akademik, maupun ketertarikan pribadi. Oleh karena itu, pendekatan *Sequential Recommender System* (SRS) berkembang untuk memodelkan urutan interaksi pengguna secara temporal [17]. Salah satu model yang banyak digunakan adalah *Self-Attentive Sequential Recommendation* (SASRec), yaitu model berbasis arsitektur transformer yang memanfaatkan mekanisme *self-attention* untuk mempelajari pola perilaku pengguna dari urutan interaksi sebelumnya [18]. Dibandingkan pendekatan tradisional, SASRec mampu menangkap dependensi jangka panjang dan perubahan preferensi pengguna secara lebih adaptif.

Akan tetapi, model tersebut umumnya masih bergantung pada representasi item berbasis ID [18]. Representasi item sebenarnya dapat ditingkatkan melalui penggunaan representasi berbasis modalitas, seperti *text embedding*. Namun, teknik *embedding* konvensional seperti Word2Vec masih memiliki keterbatasan dalam menyediakan informasi kontekstual yang mendalam [19], sehingga membatasi kemampuan model dalam menghasilkan rekomendasi yang beragam dan *serendipitous*. Untuk mengatasi masalah tersebut, *Large Language Models* (LLMs) menawarkan solusi untuk menghasilkan *embedding* yang lebih kontekstual dan mampu menangkap hubungan semantik yang lebih dalam pada data teks [20].

Beberapa penelitian sebelumnya menunjukkan bahwa penggunaan *embedding* berbasis LLM dapat meningkatkan performa sistem rekomendasi dibandingkan pendekatan *embedding* konvensional [21], [22], [23], [24]. Namun, sebagian besar penelitian masih berfokus pada peningkatan akurasi rekomendasi tanpa mengevaluasi secara mendalam aspek *serendipity* dan *trade-off* antara relevansi serta kebaruan rekomendasi. Selain itu, studi yang membandingkan berbagai jenis *embedding* modern, mulai dari Word2Vec hingga LLM terbaru, dalam konteks sistem rekomendasi sekuensial pada domain perpustakaan akademik masih relatif terbatas.

Berdasarkan permasalahan tersebut, penelitian ini mengusulkan sistem rekomendasi buku sekuensial berbasis SASRec yang diintegrasikan dengan berbagai jenis *embedding*, mulai dari Word2Vec hingga *embedding* berbasis *Large Language Model* modern. Untuk meningkatkan keseimbangan antara relevansi dan *serendipity*, penelitian ini juga menerapkan metode *Serendipity-Oriented Greedy* (SOG) *re-ranking* [25]. Penelitian difokuskan pada analisis komparatif performa *embedding* dalam menghasilkan rekomendasi yang relevan dan *serendipitous*,

termasuk pada skenario *warm-start* dan *cold-start*. Evaluasi dilakukan menggunakan metrik *Hit Ratio* (HR@K), *Normalized Discounted Cumulative Gain* (NDCG@K), dan *unSerendipity*.

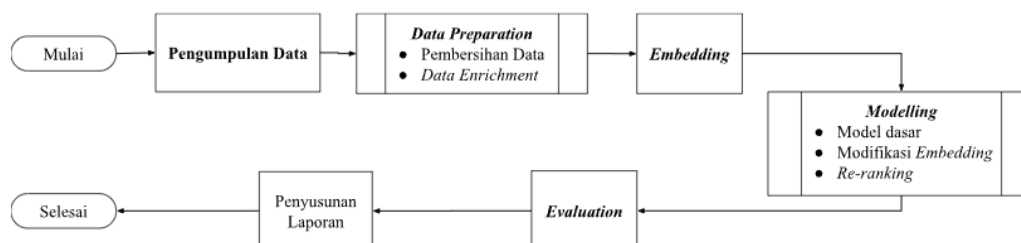
Kontribusi utama penelitian ini meliputi: (1) evaluasi komparatif *embedding* konvensional dan *embedding* berbasis LLM modern pada sistem rekomendasi sekuensial, (2) analisis pengaruh metode SOG *re-ranking* terhadap *trade-off* antara relevansi dan *serendipity*, serta (3) analisis performa model pada skenario *cold-start* dalam domain rekomendasi buku perpustakaan akademik. Diharapkan hasil penelitian ini dapat memberikan kontribusi terhadap pengembangan sistem rekomendasi perpustakaan yang lebih adaptif, eksploratif, dan mampu meningkatkan pengalaman literasi pengguna.

Meskipun penelitian ini memberikan beberapa kontribusi, terdapat sejumlah keterbatasan yang perlu diperhatikan. Pertama, evaluasi hanya dilakukan pada satu dataset perpustakaan akademik sehingga generalisasi terhadap domain lain masih terbatas. Kedua, pendekatan yang diusulkan masih bergantung pada *embedding* statis yang telah diprekomputasi, sehingga adaptabilitas terhadap preferensi pengguna yang dinamis menjadi terbatas. Ketiga, model sequential tetap bergantung pada riwayat interaksi pengguna, yang menyebabkan penurunan performa pada kondisi *cold-start*. Terakhir, evaluasi *serendipity* hanya menggunakan satu metrik, yaitu *unSerendipity*, sehingga belum sepenuhnya merepresentasikan persepsi subjektif pengguna.

## 2. METODOLOGI PENELITIAN

### 2.1. Gambaran Umum

Penelitian ini mengembangkan sistem rekomendasi buku sekuensial berbasis *Self-Attentive Sequential Recommendation* (SASRec) yang diintegrasikan dengan berbagai metode *embedding* semantik, mulai dari Word2Vec hingga *embedding* berbasis *Large Language Model* (LLM). Selain itu, penelitian juga menerapkan metode *Serendipity-Oriented Greedy* (SOG) *re-ranking* untuk meningkatkan keseimbangan antara relevansi dan *serendipity* pada hasil rekomendasi. Alur penelitian terdiri atas tahap: pengumpulan data, persiapan data, pembentukan *embedding*, pemodelan sistem rekomendasi, proses *re-ranking*, dan evaluasi performa model. Ilustrasi kerangka penelitian dapat dilihat pada Gambar 1.



Gambar 1. Tahapan Penelitian

## 2.2. Dataset

Data yang digunakan dalam penelitian ini dikumpulkan dari sistem manajemen perpustakaan pada periode 8 September 2023 hingga 18 November 2025. Dataset terdiri atas data interaksi pengguna berupa riwayat peminjaman buku serta metadata buku seperti judul, penulis, kategori, bahasa, dan deskripsi buku. Data buku awal yang berhasil dikumpulkan terdiri atas 55.160 baris data, sedangkan data riwayat peminjaman mencakup 5.445 interaksi dari 1.798 pengguna.

Untuk meningkatkan kualitas representasi teks, metadata yang tidak lengkap diperkaya menggunakan Google Books API. Penelitian ini hanya menggunakan koleksi buku berbahasa Indonesia dan Inggris serta mengecualikan dokumen nonbuku seperti skripsi, tesis, dan jurnal. Tahap persiapan data diawali dengan proses pembersihan data untuk menghapus duplikasi, memperbaiki format data, dan menangani nilai kosong. Setelah proses pembersihan, jumlah item buku yang digunakan dalam penelitian menjadi 14.502 item.

Selanjutnya dilakukan proses pembentukan sekuen dengan menyusun riwayat interaksi setiap pengguna berdasarkan urutan waktu peminjaman. Setiap urutan interaksi direpresentasikan sebagai sekuens item untuk memprediksi item berikutnya. Untuk menjaga konsistensi panjang input model, seluruh sekuens diproses menggunakan teknik padding.

Dataset kemudian dibagi menggunakan skema *Leave-One-Out* (LOO) [26], yaitu satu interaksi terakhir digunakan sebagai data uji dan satu interaksi sebelumnya sebagai data validasi, sedangkan interaksi lainnya digunakan sebagai data pelatihan. Selain itu, data uji dibagi menjadi skenario *warm-start* dan *cold-start* untuk mengevaluasi kemampuan model dalam menangani item yang telah maupun belum memiliki cukup interaksi sebelumnya. Pengguna *cold-start* ditentukan dari presentil 0.2 dari distribusi Panjang sekuens, sedangkan item *cold-start* adalah item yang belum pernah muncul sama sekali pada *training set*. *Cold-start test set* adalah gabungan dari set item yang ada di *cold-start* dan berada pada user *cold-start*.

Tabel 1. Statistik Dataset

Statistik	Nilai
Jumlah pengguna (users)	1.798
Jumlah item buku	14.502
Jumlah interaksi	5.445
Rata-rata interaksi per pengguna	2,98
Rata-rata panjang sekuens	2.99
Panjang maksimal sekuens	86
<i>Sparsity</i> dataset	99,98%
Periode data	Sept 2023 – Nov 2025
Bahasa buku	Indonesia & Inggris

## 2.3. Embedding

Dalam proses *embedding*, *Input* teks akan berupa string yang berisi gabungan data dengan skema sebagai berikut:

Judul: [judul]. Penulis: [penulis]. Deskripsi: [deskripsi], Topik: [topik], Kategori: [kategori], Bahasa: [bahasa].

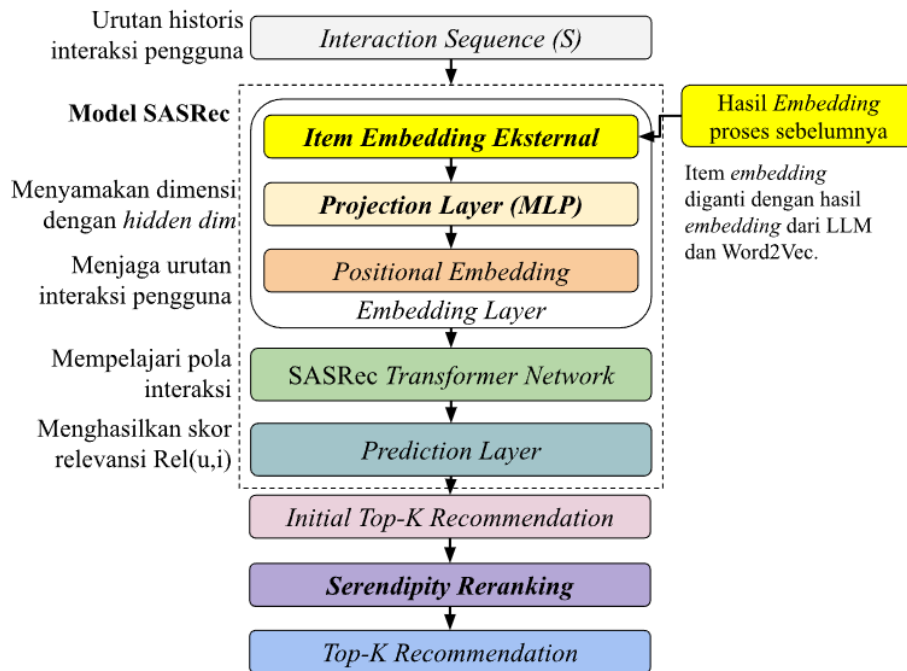
Metode *embedding* konvensional yang digunakan adalah Word2Vec, sedangkan *embedding* berbasis LLM terdiri atas beberapa model modern, yaitu bert-base-multilingual-uncased, text-embedding-3-small, Gemini-embedding-001, dan Qwen3-Embedding-0.6B.

Word2Vec menghasilkan representasi vektor statis berdasarkan kemunculan kata dalam konteks lokal, sedangkan *embedding* berbasis LLM menghasilkan *contextual embedding* yang mampu menangkap hubungan semantik dan makna konseptual teks secara lebih mendalam. Seluruh *embedding* kemudian digunakan sebagai representasi item pada model rekomendasi sekuensial.

#### 2.4. Eksperimen Model

Model utama yang digunakan dalam penelitian ini adalah SASRec (*Self-Attentive Sequential Recommendation*), yaitu model rekomendasi sekuensial berbasis arsitektur *transformer* yang memanfaatkan mekanisme *self-attention* untuk mempelajari pola interaksi pengguna secara temporal. Model ini memprediksi item berikutnya berdasarkan urutan interaksi sebelumnya.

Pada SASRec, setiap item dalam sekuens direpresentasikan sebagai *embedding* vektor. Pada penelitian ini, *embedding* item berbasis ID digantikan dengan *semantic embedding* yang telah diprekomputasi melalui kerangka *two-stage pipeline* [27]. Integrasi *embedding* dilakukan pada lapisan representasi item agar model dapat memanfaatkan informasi semantik buku secara lebih kontekstual dalam proses pembelajaran preferensi pengguna. Untuk menyelaraskan dimensi *embedding* dengan dimensi laten (*hidden dim*) model, digunakan *projection layer* berbasis *multi-layer perceptron* (MLP) [22]. *Embedding* yang telah diproyeksikan kemudian dikombinasikan dengan *positional encoding* dan diproses melalui *self-attention layer* untuk nantinya menghasilkan prediksi. Model dilatih menggunakan *binary cross-entropy loss* dan dioptimalkan menggunakan AdamW. Output model berupa daftar item kandidat yang telah diurutkan berdasarkan skor relevansi prediksi. Ilustrasi model yang digunakan dapat dilihat pada Gambar 2.



Gambar 2 Ilustrasi Model

Proses pelatihan model dilakukan menggunakan skema *next-item prediction*. Model mempelajari probabilitas kemunculan item berikutnya berdasarkan urutan interaksi pengguna sebelumnya. Konfigurasi *hyperparameter* terbaik diperoleh melalui proses *grid search* terhadap beberapa parameter seperti *learning rate*, ukuran *embedding*, jumlah *attention block*, dan jumlah lapisan linear (*num layer*).

### 2.5. Re-ranking

Untuk meningkatkan aspek *serendipity* pada hasil rekomendasi, penelitian ini menerapkan metode *Serendipity-Oriented Greedy (SOG) re-ranking*. Metode ini bekerja dengan menyusun ulang daftar rekomendasi berdasarkan kombinasi antara skor relevansi dan skor kebaruan item. Secara matematis dapat dilihat pada Persamaan (1).

$$Score_{uiB} = \alpha_{rel} \cdot \hat{r}_{ui} + \alpha_{div} \cdot div_{iB} + \alpha_{prof} \cdot prof_{ui} + \alpha_{unpop} \cdot unpop_i \quad (1)$$

Skor akhir item ( $Score_{uiB}$ ) dihitung menggunakan pembobotan ( $\alpha$ ) antara relevansi ( $\hat{r}_{ui}$ ) hasil prediksi model, tingkat ketidaksamaan semantik item terhadap riwayat interaksi pengguna ( $prof_{ui}$ ), tingkat ketidaksamaan item dengan keseluruhan koleksi ( $div_{iB}$ ), dan ( $unpop_i$ ) tingkat ketidakpopuleran item. Dengan pendekatan ini, sistem tidak hanya merekomendasikan item yang relevan, tetapi juga item yang memiliki unsur kebaruan dan keberagaman. Proses perhitungan kemiripan menggunakan *cosine similarity* seperti pada Persamaan (2).

$$CosSim(A, B) = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}} \quad (2)$$

Penelitian menguji beberapa skenario pembobotan untuk menganalisis trade-off antara akurasi dan serendipity. Skenario pembobotan dapat dilihat pada Tabel 2. Pembobotan dengan prioritas relevansi cenderung meningkatkan performa HR@K dan NDCG@K, sedangkan pembobotan yang lebih tinggi pada aspek kebaruan meningkatkan nilai unSerendipity.

Tabel 2. Tabel Bobot SOG

Alias	Rel	Div	Prof	UnPop
Weight_1	0.7	0.1	0.1	0.1
Weight_2	0.4	0.2	0.2	0.2
Weight_3	0.1	0.2	0.4	0.3
Weight_4	0.1	0.3	0.3	0.3

## 2.6. Evaluasi

Evaluasi dilakukan untuk mengukur performa model dalam menghasilkan rekomendasi yang relevan dan *serendipitous*. Penelitian menggunakan tiga metrik utama, yaitu *Hit Ratio* (HR@K), *Normalized Discounted Cumulative Gain* (NDCG@K), dan *unSerendipity@K* (unSer@K).

HR@K digunakan untuk mengukur kemampuan model dalam merekomendasikan item relevan pada daftar Top-K rekomendasi, sedangkan NDCG@K mengukur kualitas peringkat rekomendasi dengan mempertimbangkan posisi item relevan.

Sementara itu, unSerendipity@K digunakan untuk merepresentasikan *novelty* melalui rendahnya kemiripan semantik antara item dan riwayat pengguna, *unexpectedness* direpresentasikan oleh posisi item yang berada di luar pola interaksi historis pengguna, serta *relevance* dari penghitungan yang berdasarkan dengan preferensi pengguna. Berikut rumus metrik *unSerendipity* yang digunakan [28], [29]:

$$Unserendipity_u@K = \sum_{u \in U} \frac{1}{|U||H_u|} \sum_{h \in H_u} \sum_{j \in R_{u,K}} \frac{CosSim(i, h)}{K} \quad (3)$$

Nilai *UnSerendipity@K* digunakan untuk mengukur tingkat kemiripan semantik antara item rekomendasi dengan riwayat interaksi pengguna sebelumnya. Pada persamaan tersebut,  $U$  merupakan himpunan seluruh pengguna,  $H_u$  adalah himpunan buku yang pernah disukai atau diinteraksikan oleh pengguna  $u$ , sedangkan  $R_{u,K}$  merupakan daftar rekomendasi *top-K* yang dihasilkan untuk pengguna  $u$ . Fungsi  $CosSim(i, h)$  menyatakan nilai *cosine similarity* antara item rekomendasi  $i$  dan item historis  $h$ . Nilai *cosine similarity* dihitung menggunakan Persamaan (2).

Melalui pendekatan ini, semakin tinggi *unSerendipity*, semakin rendah *serendipity*, begitu juga sebaliknya. Metrik ini dipilih karena mampu mengukur tingkat ketidakterdugaan rekomendasi berdasarkan *semantic distance* antar item, sehingga

sesuai untuk mengevaluasi apakah embedding LLM maupun Word2Vec mampu menghasilkan rekomendasi yang lebih eksploratif namun tetap relevan.

Evaluasi dilakukan pada tiga skenario pengujian, yaitu keseluruhan data (*all*), *warm-start*, dan *cold-start*. Selain itu, dilakukan uji statistik menggunakan *paired t-test* untuk mengukur signifikansi perbedaan performa antar model embedding yang diuji. Setelah pengujian, dilakukan validasi signifikansi statistik menggunakan *paired t-test* dengan ambang batas  $p < 0,05$ .

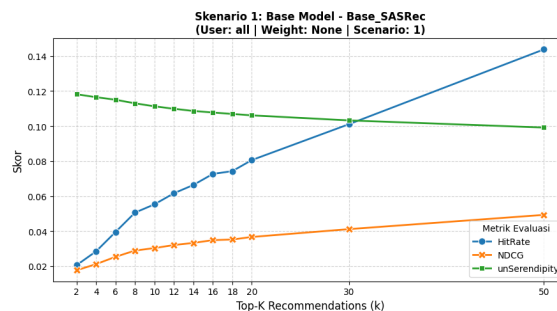
### 3. HASIL DAN PEMBAHASAN

#### 3.1. Analisis Dataset

Pada analisis panjang sekuen, ditemukan bahwa rata-rata panjang sekuen dari data yaitu sebanyak 3 dengan nilai persentil 95 adalah 9 dan nilai persentil 99 adalah 21. Distribusi ini menunjukkan adanya ketimpangan data yang sangat tinggi (*high sparsity*) pada riwayat peminjaman buku yang ada di perpustakaan. Melalui distribusi ini, panjang maksimal *padding* sekuen ditetapkan di angka 10 untuk menjaga relevansi model dan menghemat biaya komputasi.

#### 3.2. Perbandingan Varian Model

Sebagai tahap awal, penelitian mengevaluasi performa model dasar SASRec tanpa integrasi *embedding* eksternal. Model baseline digunakan sebagai acuan untuk mengukur pengaruh integrasi *embedding* semantik terhadap kualitas rekomendasi. Hasil evaluasi menunjukkan bahwa model dasar bekerja cukup baik dengan peningkatan nilai NDCG@K HR@K seiring bertambahnya nilai K. Pada K=10, HR@10, NDCG@10, dan unSer@10 berada di nilai 0.055, 0.030, dan 0.111. Visualisasi hasil evaluasi dapat dilihat pada Gambar 3.



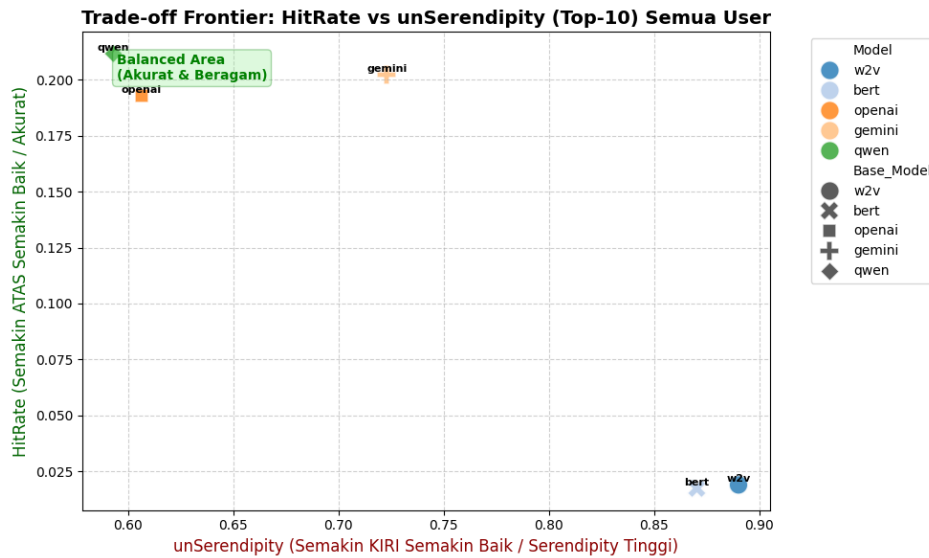
Gambar 3. Performa Model Dasar

Namun, pada model dasar ini, nilai unSer@K tercatat sangat rendah, yaitu berkisar di angka 0.1. Hal ini mengindikasikan rekomendasi yang dihasilkan model dasar cenderung mengarah pada item populer dengan tingkat kemiripan tinggi terhadap histori pengguna. Kondisi ini menunjukkan bahwa pendekatan berbasis interaksi saja belum cukup untuk memahami hubungan semantik antar buku secara mendalam. Kondisi ini mengkonfirmasi bahwa pendekatan berbasis interaksi sekuensial murni tanpa penambahan informasi semantik cenderung menghasilkan rekomendasi yang homogen dan terkonsentrasi pada item populer [30]. Setelah integrasi embedding dilakukan, terjadi perubahan performa yang cukup signifikan antar model embedding yang diuji. Hasil pengujian dapat dilihat pada Tabel 1.

Tabel 1. Tabel Hasil Evaluasi Performa Model Sebelum Re-Ranking @10

Varian	Test	HR	NDCG	unSer	%HR	%NDCG	%unSer	Selisih	Signifikan	
1	Base	Warm	0.073	0.040	0.116	-	-	-	-	-
		Cold	0.000	0.000	0.095	-	-	-	-	-
		All	0.055	0.030	0.111	-	-	-	-	-
2	w2v	all	0.031	0.009	0.890	-53.48%	-71.02%	+711.83%	-0.035	Ya
		cold	0.000	0.000	0.842	+nan%	+nan%	+799.06%	+0.000	Tidak
		warm	0.040	0.012	0.905	-53.48%	-71.02%	+689.51%	-0.047	Ya
3	Bert	all	0.035	0.014	0.870	-46.72%	-57.92%	+693.75%	-0.031	Ya
		cold	0.000	0.000	0.816	+nan%	+nan%	+771.98%	+0.000	Tidak
		warm	0.046	0.018	0.887	-46.72%	-57.92%	+673.73%	-0.041	Ya
4	Gemini	all	0.225	0.137	0.720	+240.16%	+325.06%	+556.97%	+0.159	Ya
		cold	0.000	0.000	0.655	+nan%	+nan%	+599.40%	+0.000	Tidak
		warm	0.296	0.179	0.741	+240.16%	+325.06%	+546.11%	+0.209	Ya
5	Openai	all	0.218	0.140	0.602	+229.03%	+336.22%	+449.32%	+0.152	Ya
		cold	0.004	0.001	0.525	+inf%	+inf%	+460.65%	+0.004	Ya
		warm	0.285	0.184	0.626	+227.63%	+335.44%	+446.42%	+0.198	Ya
6	Qwen	all	0.239	0.153	0.588	+260.24%	+375.28%	+436.63%	+0.172	Ya
		cold	0.006	0.002	0.510	+inf%	+inf%	+444.67%	+0.006	Ya
		warm	0.312	0.200	0.613	+258.25%	+374.04%	+434.57%	+0.225	Ya

Hasil konfigurasi terbaik tiap model pada pengujian pada @10 (Top-10 rekomendasi) menunjukkan bahwa *embedding* berbasis LLM modern secara konsisten mengungguli *embedding* tradisional seperti Word2Vec dan BERT. Pada model Word2Vec, konfigurasi terbaik menghasilkan nilai HR@10 sebesar 0,0932, NDCG@10 sebesar 0,0536, dan unSer@10 sebesar 0,7303. Sebaliknya, model Qwen menghasilkan performa yang jauh lebih tinggi dengan HR@10 mencapai 0,2433 dan NDCG@10 sebesar 0,1778, serta nilai unSer@10 sebesar 0,6061. Jika dibandingkan dengan baseline SASRec, integrasi embedding Qwen meningkatkan performa secara signifikan dengan kenaikan HitRate sebesar 282,9% dan NDCG sebesar 387,8%. Hasil tersebut menunjukkan bahwa *embedding* berbasis LLM mampu memberikan representasi semantik yang lebih efektif dibandingkan embedding statis konvensional.



Gambar 4 Trade-off pada @10 sebelum re-ranking

Selain Qwen, model Gemini dan OpenAI juga menunjukkan peningkatan performa yang konsisten. Pada visualisasi trade-off antara relevansi dan serendipity pada Gambar 4, model Qwen dan OpenAI menempati posisi paling optimal dengan nilai HR@10 di atas 0,20 dan nilai unSer@10 berada pada kisaran 0,58–0,62. Sebaliknya, embedding tradisional seperti Word2Vec, BERT, dan MiniLM terkonsentrasi pada area dengan HR@10 rendah (<0,05) dan unSer@10 tinggi (>0,85), yang menunjukkan bahwa rekomendasi cenderung kurang akurat dan tidak stabil secara semantik.

Perbedaan performa yang signifikan antara embedding berbasis LLM modern (Qwen, Gemini, OpenAI) dengan embedding tradisional (Word2Vec, BERT) merupakan temuan yang konsisten dengan penelitian yang dilakukan oleh He (2025) [23]. Model yang hanya memanfaatkan representasi tekstual memang mampu meningkatkan generalisasi, namun gagal menangkap kolerasi antar item dan pola preferensi. Hal ini juga menjelaskan mengapa Word2Vec dan BERT justru menurunkan performa HR dan NDCG dibanding baseline dan meningkatkan unSerendipity secara drastis. Embedding statis seperti Word2Vec tidak mampu menangkap nuansa semantik kontekstual yang dibutuhkan untuk memahami preferensi pengguna secara mendalam.

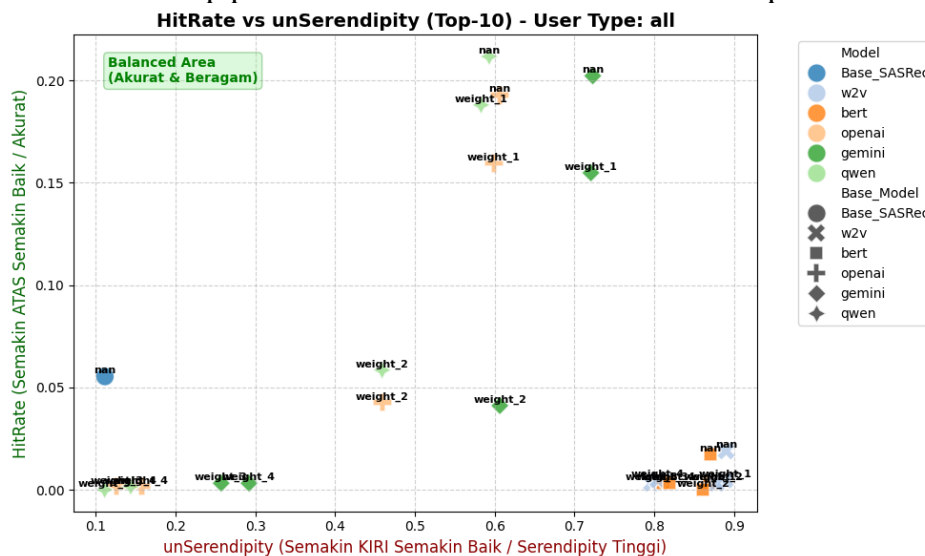
Pada skenario *warm-start*, integrasi embedding berbasis LLM menunjukkan peningkatan performa yang sangat signifikan dibandingkan baseline SASRec. Model Qwen mencatatkan peningkatan HR@10 sebesar +280,0% dan peningkatan NDCG@10 sebesar +386,1%. Harrier dan Gemini juga menunjukkan performa yang kompetitif dengan peningkatan HR@10 masing-masing sebesar +277,1% dan +265,7%. Sebaliknya, embedding tradisional seperti Word2Vec dan BERT mengalami penurunan performa cukup besar. Word2Vec mengalami penurunan HR@10 sebesar -65,7% dan NDCG@10 sebesar -77,2%, sedangkan BERT mengalami penurunan HR@10 sebesar -68,6% dan NDCG@10 sebesar -69,0%.

Pada skenario *cold-start*, seluruh model mengalami penurunan performa relevansi akibat minimnya data interaksi. Degradasi relevansi pada skenario ini mencerminkan batasan penelitian yang telah banyak didokumentasikan, salah satunya pada penelitian Du [31], [32]. Namun demikian, embedding berbasis LLM modern masih mampu mempertahankan nilai  $unSer@10$  pada kisaran menengah sebesar 0,50–0,65. Sebaliknya, model tradisional menghasilkan nilai  $unSer@10$  yang sangat tinggi dan tidak stabil, yang menunjukkan bahwa rekomendasi menjadi cenderung acak. Hal mengindikasikan bahwa *embedding* LLM sangat bermanfaat justru ketika data interaksi rendah, karena pemahaman semantiknya mendukung pemodelan dan pembelajaran representasi yang lebih baik [33]. Ini menjelaskan mengapa dalam kondisi *cold-start* sekalipun, *embedding* LLM masih unggul dari sisi kualitas representasi semantik [34]. Strategi ini dapat dipertimbangkan untuk penelitian lanjutan guna meningkatkan performa pada skenario *cold-start*.

Hasil pengujian *paired t-test* menunjukkan bahwa embedding konvensional dan LLM sederhana cenderung menghasilkan selisih performa negatif terhadap model dasar. Sebaliknya, embedding berbasis LLM modern menunjukkan peningkatan performa yang signifikan secara statistik pada hampir seluruh skenario evaluasi. Pada skenario *warm-start*, embedding LLM berperan sebagai *performance amplifier* yang secara konsisten meningkatkan kualitas rekomendasi. Sementara itu, pada skenario *cold-start*, sebagian besar model menghasilkan nilai p-value NaN akibat nilai HitRate yang konstan mendekati nol. Kondisi ini menunjukkan bahwa keterbatasan data interaksi masih menjadi tantangan utama dalam sistem rekomendasi sekuensial.

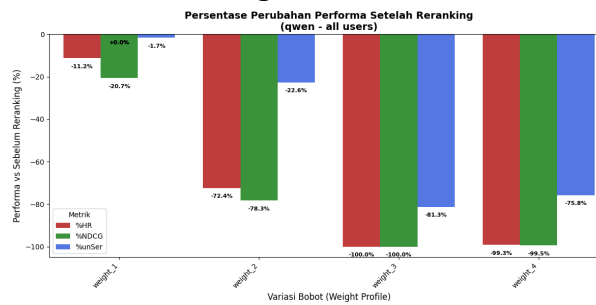
### 3.3. Re-Ranking

Tahap berikutnya adalah penerapan *Serendipity-Oriented Greedy (SOG) re-ranking* untuk menyeimbangkan relevansi dan serendipity rekomendasi. Hasil eksperimen menunjukkan bahwa perubahan bobot pada fungsi *re-ranking* memberikan dampak yang berbeda terhadap performa model. Hasil divisualisasikan pada Gambar 5.

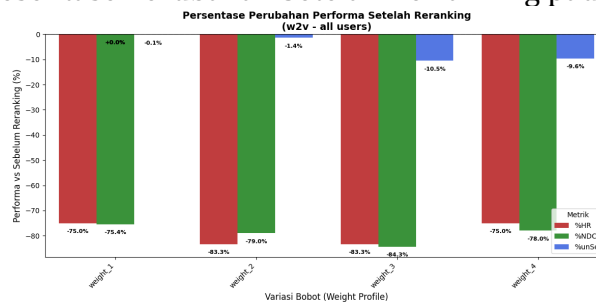


Gambar 5 Visualisasi trade-off setelah re-ranking pada all test set

Dapat ditinjau pada Gambar 5, model berbasis embedding LLM modern seperti Qwen, Gemini, dan OpenAI menunjukkan performa paling baik dibandingkan model tradisional. Pada skenario *weight\_1*, model-model tersebut berhasil menempati area paling dekat dengan *balanced area*, dengan nilai HitRate sekitar 0,15–0,21 dan nilai unSerendipity pada kisaran 0,58–0,72. Pada model Qwen, skenario *weight\_1* masih mampu mempertahankan performa relatif stabil dengan penurunan HR@10 sebesar –11,3% dan NDCG@10 sebesar –20,7%, sementara nilai unSerendipity meningkat sebesar +2,1% (Gambar 6). Kondisi ini menunjukkan bahwa pembobotan ringan masih mampu meningkatkan keberagaman rekomendasi tanpa merusak performa relevansi secara signifikan.



Gambar 6 Presentase Perubahan Setelah Re-ranking pada Model Qwen



Gambar 7 Presentase Perubahan Setelah Re-ranking pada Model Word2Vec

Namun, ketika pembobotan mulai bergeser ke skenario yang lebih agresif seperti *weight\_2*, *weight\_3*, dan *weight\_4*, seluruh model mengalami penurunan performa relevansi yang sangat signifikan. Pada *weight\_2* terjadi perubahan yang jauh lebih drastis. HR@10 turun sebesar –72,4% dan NDCG@10 sebesar –78,3%, sedangkan unSerendipity meningkat sebesar +22,4%. Penurunan semakin ekstrem pada *weight\_3* dan *weight\_4*, di mana HR@10 dan NDCG@10 hampir mencapai –100%, sementara nilai unSerendipity meningkat hingga +81,3% dan +75,8%. Hasil ini menunjukkan bahwa peningkatan keberagaman yang terlalu agresif menyebabkan rekomendasi kehilangan relevansi terhadap preferensi pengguna.

Pola serupa juga ditemukan pada model Word2Vec (Gambar 7). Pada *weight\_1*, model mengalami penurunan HR@10 sebesar –75,0% dan NDCG@10 sebesar –75,4%, sementara peningkatan unSerendipity hanya sebesar +0,1%. Bahkan ketika pembobotan eksplorasi ditingkatkan pada *weight\_2* hingga *weight\_4*, peningkatan unSerendipity tetap relatif kecil dibandingkan model berbasis LLM modern, sedangkan performa relevansi terus mengalami degradasi hingga lebih dari –80%. Hal ini menunjukkan bahwa *embedding* tradisional tidak mampu

mempertahankan kualitas rekomendasi ketika proses *re-ranking* mulai menekankan aspek keberagaman.

Fenomena ini juga mengindikasikan bahwa optimisasi *serendipity* yang terlalu agresif menyebabkan sistem kehilangan kemampuan untuk mempertahankan relevansi rekomendasi. Temuan ini juga selaras dengan Kim et al (2025) , Khaitan (2026), dan peng (2023) yang menemukan bahwa dalam mencapai *serendipity*, ada relevansi yang perlu dikorbankan [35], [36], [37].

Secara keseluruhan, hasil analisis perubahan performa menunjukkan bahwa embedding berbasis LLM modern memiliki kemampuan lebih baik dalam menjaga keseimbangan antara relevansi dan *serendipity* dibandingkan *embedding* tradisional. Selain itu, hasil ini juga memperlihatkan bahwa efektivitas algoritma SOG sangat bergantung pada pembobotan yang digunakan. Pembobotan ringan masih mampu menghasilkan rekomendasi yang seimbang, sedangkan pembobotan agresif cenderung menghasilkan rekomendasi yang beragam namun kehilangan relevansi secara hampir total.

## KESIMPULAN

Penelitian ini mengevaluasi integrasi *embedding* berbasis *Large Language Model* (LLM) pada model SASRec untuk meningkatkan kualitas rekomendasi buku pada domain perpustakaan akademik. Hasil penelitian menunjukkan bahwa *embedding* LLM modern secara konsisten mampu meningkatkan performa dibandingkan baseline SASRec dan *embedding* tradisional seperti Word2Vec dan BERT. Model Qwen3-Embedding-0.6B menghasilkan performa terbaik dengan peningkatan HR@10 hingga 282,9% dan NDCG@10 hingga 387,8%, serta mampu menjaga keseimbangan antara relevansi dan *serendipity* rekomendasi dengan lebih baik dibandingkan model lainnya. Selain itu, *embedding* berbasis LLM juga menunjukkan ketahanan semantik yang lebih stabil pada skenario *cold-start*.

Penerapan metode *Serendipity-Oriented Greedy* (SOG) *re-ranking* menunjukkan bahwa peningkatan *serendipity* memiliki *trade-off* langsung terhadap relevansi rekomendasi. Pembobotan ringan masih mampu menghasilkan keseimbangan yang baik antara akurasi dan keberagaman, namun pembobotan yang terlalu agresif menyebabkan penurunan performa relevansi secara signifikan. Temuan ini menunjukkan bahwa kualitas *embedding* semantik memiliki peran yang lebih penting dibandingkan optimisasi *re-ranking* yang agresif dalam menghasilkan rekomendasi yang relevan sekaligus eksploratif.

Secara praktis, penelitian ini menunjukkan bahwa integrasi *embedding* LLM modern pada sistem perpustakaan akademik berpotensi membantu pengguna menemukan buku yang lebih relevan dan beragam, termasuk pada koleksi dengan interaksi yang minim. Implementasi sistem rekomendasi semacam ini dapat mendukung peningkatan pengalaman pencarian informasi, eksplorasi koleksi perpustakaan, serta minat baca pengguna melalui rekomendasi yang lebih personal dan adaptif. Selain itu, pendekatan ini juga memiliki peluang untuk diterapkan pada

sistem perpustakaan nyata berbasis katalog digital maupun layanan rekomendasi daring.

Secara keseluruhan, penelitian ini menunjukkan bahwa embedding berbasis LLM modern memiliki potensi besar untuk meningkatkan kualitas sistem rekomendasi buku sekuensial, khususnya dalam menghasilkan rekomendasi yang lebih relevan, semantik, dan seimbang antara akurasi serta serendipity. Penelitian selanjutnya dapat mengeksplorasi evaluasi berbasis pengguna (*user study*), pendekatan *adaptive re-ranking*, peluang integrasi *user profiling* maupun *generative recommendation* berbasis LLM untuk menghasilkan sistem rekomendasi yang lebih personal, interaktif, dan kontekstual.

#### DAFTAR PUSTAKA

- [1] “Perpusnas RI Mewujudkan Perpustakaan sebagai Pusat Aktivitas Masyarakat - National Library of Indonesia.” Diakses: 12 Desember 2025. [Daring]. Tersedia pada: <https://perpusnas.go.id/berita/mewujudkan-perpustakaan-sebagai-pusat-aktivitas-masyarakat>
- [2] GoodStats, “Punya Perpustakaan Tertinggi di Dunia, Minat Baca di Indonesia Masih Rendah,” GoodStats. Diakses: 8 Desember 2025. [Daring]. Tersedia pada: <https://goodstats.id/article/perpus-tertinggi-dunia-namun-minat-baca-rendah-rPe7F>
- [3] W. Zarman, “Information Overload: Clarifying the Problem,” *IJIE Indones. J. Inform. Educ.*, vol. 5, no. 2, hlm. 79–83, Des 2021, doi: 10.20961/ijie.v5i2.56922.
- [4] O. Efe Stanley, “Information Overload: Causes, Symptoms, Consequences and Solutions,” *Asian J. Inf. Sci. Technol.*, vol. 11, no. 2, hlm. 1–6, Nov 2021, doi: 10.51983/ajist-2021.11.2.2887.
- [5] L. Shahrzadi, A. Mansouri, M. Alavi, dan A. Shabani, “Causes, consequences, and strategies to deal with information overload: A scoping review,” *Int. J. Inf. Manag. Data Insights*, vol. 4, no. 2, hlm. 100261, Nov 2024, doi: 10.1016/j.jjime.2024.100261.
- [6] C. C. Aggarwal, *Recommender Systems*. Cham: Springer International Publishing, 2016. doi: 10.1007/978-3-319-29659-3.
- [7] M. Sakinah dan M. Alda, “Sistem Informasi Perpustakaan SMP Negeri 1 Pegajahan Dengan Fitur Notifikasi Untuk Monitoring Peminjaman Buku Berbasis Mobile Dengan Metode Collaborative Filtering,” *Teknol. J. Ilm. Sist. Inf.*, vol. 15, no. 2, hlm. 75–86, Jul 2025, doi: 10.26594/teknologi.v15i2.5733.
- [8] M. G. Navsih, A. Muhaimin, dan S. S. M. Wara, “Sistem Rekomendasi Menu Kantin Menggunakan Lifespan-Aware Association Rule Mining Dengan Hybrid Apriori Dan FP-Growth,” *Teknol. J. Ilm. Sist. Inf.*, vol. 16, no. 1, hlm. 13–20, Feb 2026, doi: 10.26594/teknologi.v16i1.6143.
- [9] H. Asy’ari, M. Y. Urrochman, dan A. Ro’uf, “Sistem Rekomendasi Kurikulum Adaptif Berbasis Profil Lulusan Menggunakan Metode Hybrid Filtering,” *Teknol. J. Ilm. Sist. Inf.*, vol. 15, no. 2, hlm. 111–119, Sep 2025, doi: 10.26594/teknologi.v15i2.5795.
- [10] A. Zainurrohman, “Content-Based Recommender System Using NLP,” MLearning.ai. Diakses: 5 November 2025. [Daring]. Tersedia pada: <https://arif-zai-nur-rohman.medium.com/content-based-recommender-system-using-nlp-445ebb777c7a>
- [11] X. Wu, “Review of collaborative filtering recommendation systems,” *Appl. Comput. Eng.*, vol. 43, no. 1, hlm. 76–82, Feb 2024, doi: 10.54254/2755-2721/43/20230811.
- [12] T. M. A. U. Gunathilaka, P. D. Manage, J. Zhang, Y. Li, dan W. Kelly, “Addressing sparse data challenges in recommendation systems: A systematic review of rating estimation

- using sparse rating data and profile enrichment techniques,” *Intell. Syst. Appl.*, vol. 25, hlm. 200474, Mar 2025, doi: 10.1016/j.iswa.2024.200474.
- [13] Q. M. Areeb *dkk.*, “Filter Bubbles in Recommender Systems: Fact or Fallacy -- A Systematic Review,” 2 Juli 2023, *arXiv*: arXiv:2307.01221. doi: 10.48550/arXiv.2307.01221.
- [14] L. Lv, K. Q. Kang, dan G. (Gus) Liu, “Prick ‘filter bubbles’ by enhancing consumers’ novelty-seeking: The role of personalized recommendations of unmentionable products,” *Psychol. Mark.*, vol. 41, no. 10, hlm. 2355–2367, 2024, doi: 10.1002/mar.22057.
- [15] L. Chen, Y. Yang, N. Wang, K. Yang, dan Q. Yuan, “How Serendipity Improves User Satisfaction with Recommendations? A Large-Scale User Evaluation,” dalam *The World Wide Web Conference*, dalam WWW ’19. New York, NY, USA: Association for Computing Machinery, Mei 2019, hlm. 240–250. doi: 10.1145/3308558.3313469.
- [16] D. Kotkov, A. Medlar, T. Kask, dan D. Glowacka, “The Dark Matter of Serendipity in Recommender Systems,” dalam *Proceedings of the 2024 ACM SIGIR Conference on Human Information Interaction and Retrieval*, Sheffield United Kingdom: ACM, Mar 2024, hlm. 108–118. doi: 10.1145/3627508.3638342.
- [17] T. F. Boka, Z. Niu, dan R. B. Neupane, “A survey of sequential recommendation systems: Techniques, evaluation, and future directions,” *Inf. Syst.*, vol. 125, hlm. 102427, Nov 2024, doi: 10.1016/j.is.2024.102427.
- [18] W.-C. Kang dan J. McAuley, “Self-Attentive Sequential Recommendation,” dalam *2018 IEEE International Conference on Data Mining (ICDM)*, Nov 2018, hlm. 197–206. doi: 10.1109/ICDM.2018.00035.
- [19] D. S. Asudani, N. K. Nagwani, dan P. Singh, “Impact of word embedding models on text analytics in deep learning environment: a review,” *Artif. Intell. Rev.*, hlm. 1–81, Feb 2023, doi: 10.1007/s10462-023-10419-1.
- [20] S. Deb dan A. K. Chanda, “Comparative analysis of contextual and context-free embeddings in disaster prediction from Twitter data,” *Mach. Learn. Appl.*, vol. 7, hlm. 100253, Mar 2022, doi: 10.1016/j.mlwa.2022.100253.
- [21] A. Boz *dkk.*, “Improving Sequential Recommendations with LLMs,” *ACM Trans. Recomm. Syst.*, hlm. 3711667, Jan 2025, doi: 10.1145/3711667.
- [22] J. Hu *dkk.*, “Enhancing Sequential Recommendation via LLM-based Semantic Embedding Learning,” dalam *Companion Proceedings of the ACM Web Conference 2024*, dalam WWW ’24. New York, NY, USA: Association for Computing Machinery, Mei 2024, hlm. 103–111. doi: 10.1145/3589335.3648307.
- [23] Y. He, X. Liu, A. Zhang, Y. Ma, dan T.-S. Chua, “LLM2Rec: Large Language Models Are Powerful Embedding Models for Sequential Recommendation,” 16 Juni 2025, *arXiv*: arXiv:2506.21579. doi: 10.48550/arXiv.2506.21579.
- [24] K. Liu, “When Transformers Meet Recommenders: Integrating Self-Attentive Sequential Recommendation with Fine-Tuned LLMs,” 8 Juli 2025, *arXiv*: arXiv:2507.05733. doi: 10.48550/arXiv.2507.05733.
- [25] D. Kotkov, J. Veijalainen, dan S. Wang, “How does serendipity affect diversity in recommender systems? A serendipity-oriented greedy algorithm,” *Computing*, vol. 102, no. 2, hlm. 393–411, Feb 2020, doi: 10.1007/s00607-018-0687-5.
- [26] G. C. Cawley, “Leave-One-Out Cross-Validation Based Model Selection Criteria for Weighted LS-SVMs,” dalam *The 2006 IEEE International Joint Conference on Neural Network Proceedings*, Jul 2006, hlm. 1661–1668. doi: 10.1109/IJCNN.2006.246634.
- [27] Z. Yuan *dkk.*, “Where to Go Next for Recommender Systems? ID- vs. Modality-based Recommender Models Revisited,” 2 Mei 2023, *arXiv*: arXiv:2303.13835. doi: 10.48550/arXiv.2303.13835.

- [28] Y. Kim, S.-Y. Kim, dan B. Park, "A Serendipity Recommendation Method for Book Categories Using BERT to Strengthen the Web Service of the Book," *J. Web Eng.*, hlm. 199–216, Apr 2025, doi: 10.13052/jwe1540-9589.2422.
- [29] K. Zuva dan T. Zuva, "Diversity and Serendipity in Recommender Systems," dalam *Proceedings of the International Conference on Big Data and Internet of Thing*, dalam BDIOT '17. New York, NY, USA: Association for Computing Machinery, Des 2017, hlm. 120–124. doi: 10.1145/3175684.3175694.
- [30] S. Karimi, H. A. Rahmani, M. Naghiaei, dan L. Safari, "Provider Fairness and Beyond-Accuracy Trade-offs in Recommender Systems," 8 September 2023, *arXiv*: arXiv:2309.04250. doi: 10.48550/arXiv.2309.04250.
- [31] L. Pan, W. Pan, M. Wei, H. Yin, dan Z. Ming, "A Survey on Sequential Recommendation," 17 Desember 2024, *arXiv*: arXiv:2412.12770. doi: 10.48550/arXiv.2412.12770.
- [32] W.-W. Du, T. Udagawa, dan K. Tateno, "Not Just What, But When: Integrating Irregular Intervals to LLM for Sequential Recommendation," *arXiv.org*. Diakses: 15 Mei 2026. [Daring]. Tersedia pada: <https://arxiv.org/abs/2507.23209v1>
- [33] B.-C. Chen dan M. Slokom, "LLMDiRec: LLM-Enhanced Intent Diffusion for Sequential Recommendation," 20 Oktober 2025, *arXiv*: arXiv:2601.03259. doi: 10.48550/arXiv.2601.03259.
- [34] A. Pembek, A. Fatkulin, A. Klenitskiy, dan A. Vasilev, "Let It Go? Not Quite: Addressing Item Cold Start in Sequential Recommendations with Content-Based Initialization," 25 Juli 2025, *arXiv*: arXiv:2507.19473. doi: 10.48550/arXiv.2507.19473.
- [35] Y. Kim, S. Oh, C. Noh, E. Hong, dan S. Park, "Design of a Serendipity-Incorporated Recommender System," *Electronics*, vol. 14, no. 4, hlm. 821, Jan 2025, doi: 10.3390/electronics14040821.
- [36] S. Khaitan dan R. Shrivastava, "Developing Fairness, Accuracy, and Serendipity Objective Functions for Recommendation System and Establishing Trade-off through Multi-Objective Evolutionary Optimization," *Inf. Process. Manag.*, vol. 63, no. 4, hlm. 104604, Jun 2026, doi: 10.1016/j.ipm.2025.104604.
- [37] K. Peng, M. Raghavan, E. Pierson, J. Kleinberg, dan N. Garg, "Reconciling the accuracy-diversity trade-off in recommendations," *arXiv.org*. Diakses: 15 Mei 2026. [Daring]. Tersedia pada: <https://arxiv.org/abs/2307.15142v1>