

Sistem Rekomendasi berbasis Teks untuk Masalah *User Cold-Start* pada Rekomendasi Produk *E-Commerce*: Sebuah Tinjauan Literatur Sistematis

Text-Based Recommender Systems for User Cold-Start Problems in E-Commerce Product Recommendation: A Systematic Literature Review

¹Muhammad Yusril Helmi Setyawan*, ²Haris Saefuloh

^{1,2}Teknik Informatika, Fakultas School of Information Technology, Universitas Logistik dan Bisnis Internasional

^{1,2}Jl. Sariasih No. 54, Bandung, Jawa Barat, Indonesia

*e-mail: ¹yusrilhelmi@ulbi.ac.id, ²harissaefuloh@gmail.com

(received: 2 August 2026, revised: 25 September 2026, accepted: 26 September 2026)

Abstrak

Masalah *user cold-start* masih menjadi tantangan utama dalam rekomendasi produk *e-commerce* yang dipersonalisasi karena pengguna baru hanya memiliki sedikit atau belum memiliki riwayat interaksi, sehingga efektivitas *collaborative filtering* konvensional dan model berbasis *rating* menjadi terbatas. Informasi tekstual seperti ulasan pengguna, ulasan item, deskripsi produk, metadata, dan atribut semantik produk berpotensi menjadi sinyal awal untuk memodelkan preferensi pengguna dan karakteristik produk. Penelitian ini bertujuan mengidentifikasi sumber data tekstual, metode AI/NLP, dan keluarga model yang sesuai untuk menangani *user cold-start* pada rekomendasi produk *e-commerce*. Tinjauan literatur sistematis dilakukan mengikuti kerangka PRISMA 2020 terhadap 55 artikel terpilih dari basis data Scopus yang disaring dan diekstraksi dengan bantuan Watase, dikategorikan menjadi 31 studi inti dan 24 studi pendukung, serta dinilai menggunakan lima kriteria kualitas. Berbeda dari tinjauan *cold-start* terdahulu yang memetakan strategi mitigasi secara luas, tinjauan ini mempersempit cakupan pada bukti berbasis teks dan memasang setiap keluarga model dengan kondisi data yang menjadi prasyarat penerapannya. Hasil menunjukkan teks ulasan sebagai sumber tekstual dominan dengan 33 dari 55 studi, diikuti deskripsi produk dan metadata; studi terpetakan ke dalam tujuh keluarga model, dan 43 studi atau 78,2% melaporkan metrik evaluasi secara eksplisit. Arah model yang menonjol mencakup rekomendasi *hybrid deep sentiment-aware*, model *transformer* atau *attention* berbasis ulasan, rekomendasi berbasis SBERT/BERT, rekomendasi berbasis graf, dan rekomendasi semantik berbasis ontologi. Berdasarkan bukti terlapor, arah yang paling dapat dipertahankan adalah *recommender hybrid text-semantic*, meskipun perbedaan *dataset*, protokol evaluasi, dan metrik antar studi membuat perbandingan performa secara langsung tidak dapat dilakukan.

Kata kunci: *e-commerce*, rekomendasi produk, sistem rekomendasi berbasis teks, tinjauan literatur sistematis, *user cold-start*

Abstract

The *user cold-start* problem remains a major challenge in personalized *e-commerce* product recommendation because new users have little or no interaction history, limiting the effectiveness of conventional *collaborative filtering* and *rating-based* models. Textual information, such as user reviews, item reviews, product descriptions, metadata, and semantic product attributes, can provide valuable initial signals for modeling user preferences and product characteristics. This study aims to identify textual data sources, AI/NLP methods, and model families suitable for addressing the *user cold-start* problem in *e-commerce* product recommendation. A systematic literature review was conducted following the PRISMA 2020 framework, covering 55 selected articles from the Scopus database that were screened and extracted with the assistance of Watase. The studies were categorized into 31 core studies and 24 supporting studies and assessed using five quality criteria. Unlike previous *cold-start* reviews that broadly map mitigation strategies, this review narrows its scope to text-based evidence and links each model family to the data conditions required for its

<http://sistemasi.ftik.unisi.ac.id>

application. The results show that review text is the dominant textual data source, appearing in 33 of the 55 studies, followed by product descriptions and metadata. The studies were mapped into seven model families, and 43 studies (78.2%) explicitly reported evaluation metrics. Prominent model directions include hybrid deep sentiment-aware recommenders, review-based transformer or attention models, SBERT/BERT-based recommendation, graph-based recommendation, and ontology-based semantic recommendation. Based on the reported evidence, text-semantic hybrid recommenders represent a promising direction, although differences in datasets, evaluation protocols, and metrics across studies prevent direct performance comparisons.

Keywords: e-commerce, product recommendation, systematic literature review, text-based recommender systems, user cold-start.

1 Pendahuluan

Rekomendasi produk yang dipersonalisasi merupakan fungsi penting dalam *e-commerce* karena pengguna sering dihadapkan pada sejumlah besar produk, varian, dan alternatif yang saling bersaing. Sistem rekomendasi mengurangi kelebihan beban informasi ini dengan memperkirakan produk mana yang paling mungkin sesuai dengan kebutuhan pengguna [1]. Dalam lingkungan belanja daring, personalisasi juga mengurangi upaya yang dibutuhkan untuk menelusuri dan membandingkan produk. Peran ini menjadi semakin penting seiring bertambahnya ukuran katalog dan meningkatnya ekspektasi pengguna akan saran produk yang relevan sejak awal kunjungan [2]. Kualitas representasi pengguna dan item sangat memengaruhi kualitas rekomendasi. Representasi yang lemah membuat proses perankingan menjadi tidak stabil, terutama ketika sistem belum cukup mengamati perilaku pengguna [3].

Masalah *user cold-start* terjadi ketika pengguna baru hanya memiliki sedikit atau bahkan tidak memiliki riwayat interaksi. Dalam kondisi ini, *rating*, klik, keranjang, pembelian, dan pola penelusuran belum cukup untuk membangun profil preferensi yang andal [4]. *Collaborative filtering* dan *matrix factorization* dapat bekerja dengan baik ketika data interaksi tersedia. Keterbatasannya terlihat ketika pengguna baru belum menghasilkan interaksi yang cukup untuk pembelajaran berbasis tetangga (*neighborhood*) maupun faktor laten [5]. Metode berbasis *rating* tetap berguna sebagai *baseline* untuk mengukur perilaku rekomendasi. Namun, metode tersebut biasanya belum memadai sebagai solusi utama untuk *cold-start* pengguna baru karena masih bergantung pada sinyal interaksi historis [6].

Data tekstual dapat mengurangi keterbatasan tersebut karena informasi produk telah tersedia sebelum riwayat pengguna jangka panjang terkumpul. Deskripsi produk dapat menjelaskan karakteristik item yang belum pernah diklik atau dibeli oleh pengguna baru [7]. Ulasan pengguna juga penting karena memuat opini, sentimen, dan penilaian pada tingkat aspek yang tidak terlihat dari *rating* numerik semata. Hal ini membuat ulasan berguna untuk menurunkan sinyal preferensi dalam skenario *cold-start* [8]. Representasi semantik dapat menghubungkan teks produk dengan preferensi pengguna secara lebih langsung dibandingkan catatan interaksi yang jarang (*sparse*). Representasi berbasis SBERT, misalnya, dapat menyandikan deskripsi atau ulasan menjadi vektor semantik yang dapat dibandingkan [9].

Bukti berbasis tinjauan lain turut mendukung rasional ini. Rekomendasi *e-commerce* kontekstual dapat mengintegrasikan *embedding* tekstual dengan metadata pengguna dan item untuk mengurangi *sparsity* [22], sementara studi prediksi *rating* berbasis teks ulasan menunjukkan bahwa topik, sentimen, aspek linguistik, dan kata fitur menyediakan pengetahuan preferensi di luar *rating* numerik [23]. Rekomendasi berbasis aspek lebih jauh menunjukkan bahwa kelompok opini dan tingkat kepentingan aspek dapat mendukung rekomendasi yang dipersonalisasi dan dapat dijelaskan dari teks ulasan [24].

Model teks sekuensial dan *hybrid* juga menunjukkan bahwa informasi tekstual dapat melengkapi model rekomendasi tradisional. Representasi pada tingkat kata dari ulasan dapat menambahkan petunjuk preferensi yang tidak tersedia pada data berbasis *rating* saja [10]. Kebaruan (*novelty*) dari tinjauan literatur sistematis ini adalah pemetaan yang terfokus terhadap sumber data tekstual, metode AI/NLP, dan keluarga model untuk *user cold-start* pada rekomendasi produk *e-commerce*. Topik yang lebih luas seperti rekomendasi *multimodal*, privasi, dan *explainability* diperlakukan hanya sebagai tren pendukung ketika muncul dalam studi terpilih [11]. Tujuan penelitian ini adalah

mengidentifikasi sumber data tekstual, metode AI/NLP, serta keluarga model yang paling sesuai berdasarkan bukti terlapor, sehingga hasilnya dapat menjadi acuan dalam merancang sistem rekomendasi produk yang tetap andal untuk pengguna baru.

Tinjauan sekunder yang telah tersedia belum menjawab kebutuhan tersebut secara memadai, dan perbedaannya perlu dinyatakan secara spesifik. Yuan dan Hernandez [4] memetakan *user cold-start* sebagai persoalan tersendiri, tetapi tidak memilah kontribusi masing-masing jenis sumber tekstual terhadap pembentukan profil pengguna baru. Ray [19] menyusun taksonomi strategi mitigasi *cold-start* yang luas, namun menempatkan pendekatan berbasis teks sebagai satu kategori tunggal sehingga perbedaan antara representasi *embedding*, fitur sentimen, dan relasi pengetahuan tidak terlihat. Vahidy dkk. [5] dan Dong dkk. [17] menjelaskan sebab melemahnya model konvensional pada data jarang, tetapi tidak sampai membandingkan keluarga model berbasis teks untuk skenario pengguna baru, sedangkan Raeesi dkk. [20] berhenti pada identifikasi tren riset. Dengan demikian, celah yang tersisa tidak terletak pada ketersediaan strategi mitigasi, melainkan pada belum adanya pemetaan yang menautkan jenis sumber tekstual dengan keluarga model sekaligus kondisi data yang menjadi prasyarat penerapannya.

Atas dasar celah tersebut, tinjauan ini memberikan tiga kontribusi yang saling terkait. Pertama, tinjauan ini memetakan jenis sumber data tekstual yang digunakan pada skenario *user cold-start* beserta sebarannya dalam korpus, dan kontribusi ini dijawab melalui RQ1. Kedua, tinjauan ini menyusun peta metode AI, *machine learning*, *deep learning*, dan NLP yang bekerja di atas sumber tekstual tersebut beserta pola kombinasinya, dan kontribusi ini dijawab melalui RQ2. Ketiga, tinjauan ini menyusun perbandingan antar-keluarga model yang tidak berhenti pada capaian metrik, melainkan menautkan kelebihan, keterbatasan yang diakui studi asal, dan kondisi data yang menjadi prasyarat penerapan, dan kontribusi ini dijawab melalui RQ3. Ketiga kontribusi tersebut dilaporkan secara berurutan pada bagian Hasil dan Pembahasan.

2 Tinjauan Literatur

User cold-start umumnya ditangani dengan memanfaatkan informasi sampingan (*side information*) di luar riwayat interaksi jangka panjang pengguna. *Active learning* merupakan salah satu strategi karena berusaha memperoleh informasi preferensi secara lebih terarah ketika data pengguna masih jarang [27]. *Collaborative filtering* dan *matrix factorization* tetap menjadi fondasi umum dalam sistem rekomendasi, namun kelemahannya adalah keduanya membutuhkan interaksi pengguna-item yang cukup padat sebelum dapat menyimpulkan preferensi yang stabil [13]. Studi sekunder tentang *cold-start* menunjukkan bahwa masalah ini dapat melibatkan pengguna baru, item baru, atau domain baru. Tinjauan ini berfokus pada *cold-start* pengguna baru karena tujuan penelitiannya adalah rekomendasi produk yang dipersonalisasi bagi pengguna dengan riwayat terbatas [4]. *Baseline* berorientasi peringkat seperti RBPR berguna untuk pembandingan dalam studi *cold-start*, tetapi kurang sesuai sebagai jawaban tunggal untuk *user cold-start* berbasis teks karena masih memerlukan bukti interaksi atau *rating* [6].

Beberapa strategi non-teks atau sebagian berbasis teks juga menangani *user cold-start* dan *sparsity*. *Deep learning* berbasis *switching* menyesuaikan strategi rekomendasi untuk pengguna baru, pengguna ringan, dan pengguna berat [25]; metode kemiripan sesi memanfaatkan informasi sesi kontekstual dan temporal ketika sebuah sesi baru belum memiliki interaksi item sebelumnya [26]; *active learning* meminta pengguna baru menilai item terpilih [27]; dan *embedding* berbasis metadata dapat merepresentasikan pengguna dan item baru ketika interaksi historis tidak tersedia [28]. Kerangka industrial dan berbasis *transformer* terkini juga menangani rekomendasi *few-shot*, *zero-shot*, dan berbasis konteks melalui penyajian *cold-start* yang *model-agnostic* atau perankingan *transformer* yang diperkaya retrieval [13], [29].

Pada sisi teks, rekomendasi produk memanfaatkan sumber seperti ulasan, deskripsi, judul, kategori, tag, atribut, metadata, dan kata kunci preferensi, yang memungkinkan sistem memahami produk dan preferensi sebelum riwayat pengguna yang panjang terbentuk [7]. Sebagian pendekatan berupaya menyaring sinyal preferensi dari teks ulasan yang panjang dan berulang: SuReFAR memanfaatkan ringkasan ulasan dan *neural representation learning* agar ulasan dapat menjembatani informasi produk dengan bukti preferensi pengguna [14], sedangkan SRARS memadukan peringkasan

ulasan dengan mekanisme *attention* untuk menekan derau (noise) sebelum sinyal preferensi digunakan pada prediksi rekomendasi [37].

Fitur tekstual pada model berorientasi produk juga dapat diturunkan dari spesifikasi pelanggan, sisi ulasan, topik semantik, maupun *topic embedding*. Sebagian model menautkan teks ke penilaian eksplisit, misalnya PRUS yang memeringkat produk dengan mencocokkan spesifikasi pengguna terhadap polaritas ulasan positif dan negatif [31], sementara IGRc membangun graf pengguna-ulasan-item untuk mempelajari interaksi berbasis ulasan [32]. Pendekatan berbasis topik seperti *topic model* dan *topic embedding* memanfaatkan LDA, *word embedding*, dan informasi kontekstual untuk menangani rekomendasi yang *sparse* atau *cold-start* [33], [34], dan Context2Rec memperluas gagasan ini dengan menafsirkan teks kontekstual untuk memperkuat rekomendasi sekuensial saat profil pengguna belum stabil [16]. Ketika teks dapat dihubungkan dengan struktur relasional, sinyalnya diperkuat oleh relasi produk-pengguna, sebagaimana Talk-with-Graph yang memadukan *large language model*, analisis sentimen, *fuzzy logic*, dan *heterogeneous graph neural network* [38], serta RAKCR yang memakai jalur penalaran dan *knowledge graph* untuk merepresentasikan produk sebagai relasi semantik [18]. Studi *cross-domain* dan *multimodal* dipertahankan terutama sebagai bukti pendukung, mencakup *deep shared learning* untuk transfer preferensi antar-domain [35], *collaborative filtering* berbasis *biclustering* pada kondisi *sparse* atau *cold-start* [36], dan review-driven *hyperbolic embedding* yang mempertahankan struktur hierarkis pada rekomendasi lintas domain [15].

Beberapa studi sekunder telah mengkaji sistem rekomendasi dan *cold-start* dari sudut pandang yang berbeda. Studi-studi ini berguna sebagai latar belakang karena menjelaskan lanskap umum masalah *cold-start*, strategi mitigasi, dan tren sistem rekomendasi. Namun, sebagian besar tidak secara khusus berfokus pada *user cold-start* berbasis teks pada rekomendasi produk *e-commerce* yang dipersonalisasi. Yuan dan Hernandez [4] meninjau masalah *user cold-start* dalam sistem rekomendasi; studinya relevan karena menjelaskan *cold-start* sebagai tantangan spesifik, tetapi cakupannya lebih luas dan tidak berkonsentrasi pada sumber data tekstual. Ray [19] melakukan tinjauan literatur sistematis mengenai pendekatan dan algoritma untuk memitigasi masalah *cold-start* dan memetakan beberapa keluarga solusi, namun pembahasannya bersifat umum dan tidak menekankan rekomendasi produk berbasis teks pada *e-commerce*. Vahidy dkk. [5] membahas sistem rekomendasi dan *cold-start* sebagai isu yang lebih luas, membantu menjelaskan mengapa pendekatan konvensional melemah ketika data interaksi jarang, tetapi tidak secara spesifik membandingkan model representasi teks, model *sentiment-aware*, model berbasis *transformer*, atau model semantik berbasis graf. Raeesi dkk. [20] menggunakan algoritma *text mining* untuk mengidentifikasi tren yang sedang berkembang, namun tujuannya adalah identifikasi tren dan bukan evaluasi model untuk skenario *cold-start*. Dong dkk. [17] juga membahas sistem rekomendasi dan *cold-start*, tetapi tidak secara khusus membahas model AI/NLP berbasis teks mana yang paling sesuai untuk menyelesaikan *user cold-start* pada rekomendasi produk *e-commerce*. Ringkasan studi sekunder terkait disajikan pada Tabel 1.

Tabel 1 Studi terkait

Referensi	Fokus	Keterbatasan
[4]	Meninjau masalah <i>user cold-start</i> dalam sistem rekomendasi.	Tidak secara khusus berfokus pada sumber data tekstual dan rekomendasi produk <i>e-commerce</i> .
[19]	Meninjau pendekatan dan algoritma untuk memitigasi masalah <i>cold-start</i> .	Membahas <i>cold-start</i> secara luas dan tidak memprioritaskan model AI/NLP berbasis teks.
[5]	Membahas sistem rekomendasi dan <i>cold-start</i> .	Memberikan latar belakang umum, tetapi tidak membandingkan keluarga model berbasis teks untuk skenario pengguna baru.
[20]	Menggunakan <i>text mining</i> untuk mengidentifikasi tren yang berkembang dalam riset sistem rekomendasi.	Berfokus pada identifikasi tren, bukan efektivitas model untuk <i>user cold-start</i> .
[17]	Membahas sistem rekomendasi dan <i>cold-start</i> .	Tidak secara khusus mengevaluasi model berbasis teks untuk rekomendasi produk <i>e-commerce</i> yang dipersonalisasi.

Berdasarkan studi-studi sekunder tersebut, celah penelitian menjadi jelas. Tinjauan yang ada telah membahas *cold-start*, algoritma sistem rekomendasi, atau tren riset, tetapi masih diperlukan tinjauan yang lebih terfokus yang memetakan sumber data tekstual, metode AI/NLP, dan keluarga model efektif untuk *user cold-start* pada rekomendasi produk *e-commerce* yang dipersonalisasi. Oleh karena itu, penelitian ini memosisikan diri sebagai tinjauan sistematis yang terfokus pada sistem rekomendasi berbasis teks untuk masalah *user cold-start* pada rekomendasi produk *e-commerce*.

3 Metode Penelitian

Penyusunan tinjauan literatur sistematis ini diorganisasikan ke dalam tiga tahap utama, yaitu perencanaan tinjauan (review planning), pelaksanaan tinjauan (conducting the review), dan pelaporan tinjauan (reporting the review). Struktur ini diadaptasi dari pola penulisan tinjauan sistematis pada studi SLR sebelumnya, yang menjelaskan bagaimana pertanyaan penelitian, kata kunci pencarian, kriteria inklusi dan eksklusi, seleksi studi, penilaian kualitas, serta ekstraksi data didefinisikan sebelum hasil disintesis [4], [19]. Tujuan tahap ini adalah memastikan bahwa studi terpilih relevan dengan sistem rekomendasi berbasis teks, masalah *user cold-start*, dan rekomendasi produk yang dipersonalisasi pada *e-commerce*.

Pada tahap perencanaan, tujuan penelitian, pertanyaan penelitian, strategi pencarian, dan kriteria seleksi ditetapkan. Tahap ini penting karena keluaran awal Watase memuat studi dari beberapa domain sistem rekomendasi, sehingga cakupannya perlu dipersempit agar tinjauan tetap terfokus pada *user cold-start* berbasis teks pada rekomendasi produk *e-commerce*. Tujuan utama tinjauan ini adalah mengidentifikasi sumber data tekstual, metode AI/NLP, dan keluarga model yang efektif untuk menangani masalah *user cold-start*. Untuk mencapainya, dirumuskan tiga pertanyaan penelitian berikut:

RQ1: Jenis data tekstual apa yang digunakan untuk memodelkan preferensi pengguna dan informasi produk pada sistem rekomendasi produk yang dipersonalisasi untuk skenario *user cold-start*?

RQ2: Metode AI, *machine learning*, *deep learning*, dan NLP apa yang digunakan untuk menangani masalah *user cold-start* pada sistem rekomendasi produk berbasis teks yang dipersonalisasi?

RQ3: Metode atau model mana yang menunjukkan performa paling efektif dalam menyelesaikan masalah *user cold-start* berdasarkan metrik evaluasi dan hasil eksperimen?

RQ1 dirumuskan untuk mengidentifikasi sumber tekstual yang digunakan pada studi terpilih, seperti ulasan pengguna, ulasan item, deskripsi produk, metadata, atribut produk, kategori produk, dan kata kunci preferensi. RQ2 dirumuskan untuk memetakan metode yang digunakan dalam mengolah sumber tekstual tersebut, termasuk rekomendasi *hybrid*, rekomendasi *sentiment-aware*, representasi berbasis *embedding*, *deep learning*, model berbasis *transformer*, rekomendasi berbasis graf, dan rekomendasi berbasis ontologi. RQ3 dirumuskan untuk mengidentifikasi keluarga model yang paling sesuai; dalam tinjauan ini metrik evaluasi diperlakukan sebagai bukti pendukung, sedangkan jawaban utama RQ3 berfokus pada nama dan keluarga model. Fokus penelitian ini dirangkum menggunakan kerangka PICOC sebagaimana disajikan pada Tabel 2.

Tabel 2 Kerangka PICOC dan fokus penelitian

Elemen	Definisi
Population	Sistem rekomendasi berbasis teks, sistem rekomendasi yang dipersonalisasi, dan sistem rekomendasi produk.
Intervention	AI, <i>Machine Learning</i> , <i>Deep Learning</i> , NLP, <i>text mining</i> , analisis semantik, metode <i>embedding</i> , dan model <i>transformer</i> .
Comparison	Perbandingan antar-pendekatan rekomendasi, model AI, teknik NLP, dan metode penanganan <i>cold-start</i> .
Outcome	Metode/model yang efektif untuk menyelesaikan masalah <i>user cold-start</i> yang diukur melalui performa rekomendasi.
Context	Lingkungan rekomendasi produk <i>e-commerce</i> .

Setelah tujuan dan pertanyaan penelitian ditetapkan, kata kunci pencarian diturunkan dari elemen PICOC. Population mencakup sistem rekomendasi berbasis teks, sistem rekomendasi yang dipersonalisasi, dan sistem rekomendasi produk. Intervention mencakup AI, *machine learning*, *deep learning*, NLP, analisis semantik, metode *embedding*, dan model *transformer*. Context dibatasi pada *e-commerce*, belanja daring, ritel daring, dan lingkungan rekomendasi produk. Istilah pencarian yang digunakan pada proses penyaringan Watase antara lain *personalized recommendation e-commerce*, *personalized recommendation product*, *user cold start problem recommendation*, *cold start recommendation system*, *product recommender system*, *hybrid personalized recommender system*, *recommender system e-commerce*, dan *product*. Istilah tersebut digunakan untuk menarik rekaman dari Scopus, kemudian disempurnakan berdasarkan tahun, ketersediaan dokumen, tingkat publikasi, dan relevansi abstrak. Strategi pencarian dapat dinyatakan sebagai berikut:

("personalized recommendation" OR "personalized recommender system" OR "product recommender system" OR "recommender system") AND ("e-commerce" OR "electronic commerce" OR "online shopping" OR "online retail" OR "product recommendation") AND ("cold start" OR "cold-start" OR "user cold-start" OR "data sparsity") AND ("text" OR "review" OR "product description" OR "metadata" OR "NLP" OR "sentiment" OR "embedding" OR "BERT" OR "transformer")

Ringkasan strategi pencarian yang mencakup basis data, fokus pencarian, kata kunci utama dan pendukung, rentang tahun, bahasa, serta jenis dokumen disajikan pada Tabel 3.

Tabel 3 Strategi pencarian

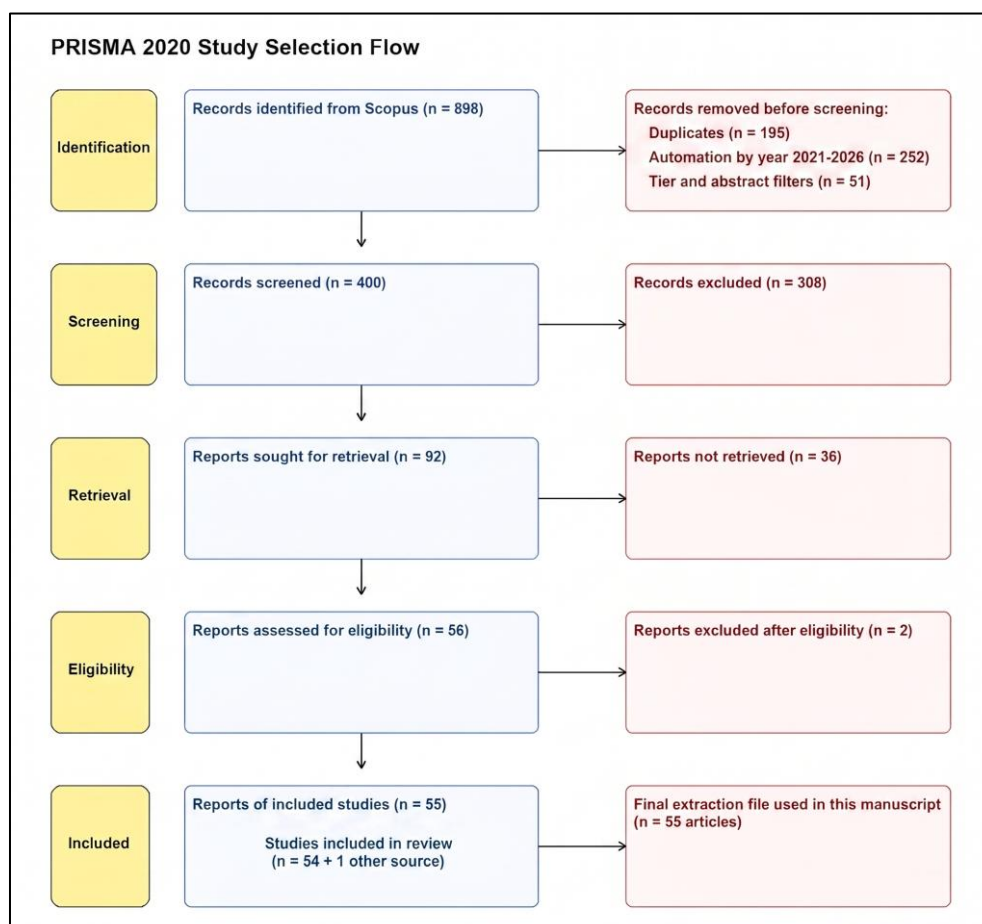
Aspek	Keterangan
Basis data	Scopus
Fokus pencarian	Rekomendasi produk yang dipersonalisasi, rekomendasi <i>e-commerce</i> , <i>user cold-start</i> , rekomendasi berbasis teks
Kata kunci utama	<i>personalized recommendation</i> , <i>product recommender system</i> , <i>text-based recommender system e-commerce</i> , <i>user cold start recommendation system</i>
Kata kunci pendukung	<i>review text</i> , <i>product description</i> , <i>metadata</i> , <i>NLP</i> , <i>sentiment analysis</i> , <i>embedding</i> , <i>BERT</i> , <i>transformer</i>
Tahun publikasi	2021-2026
Bahasa	Inggris
Jenis dokumen	Artikel penelitian dan artikel review yang relevan

Scopus dipilih sebagai basis data tunggal atas tiga pertimbangan yang sekaligus menjadi batas cakupan tinjauan ini. Pertama, Scopus mengindeks penerbit yang paling banyak memuat riset sistem rekomendasi, seperti Elsevier, Springer, IEEE, ACM, MDPI, dan Wiley, sehingga cakupan lintas-penerbit dapat dicapai melalui satu sintaks pencarian yang konsisten. Kedua, metadata Scopus menyediakan bidang terstruktur berupa abstrak, kata kunci, tahun terbit, jenis dokumen, dan afiliasi yang dibutuhkan oleh proses penyaringan serta klasifikasi berbantuan perkakas pada tinjauan ini. Ketiga, penggunaan satu basis data dengan satu sintaks membuat proses pencarian dapat diulang oleh peneliti lain tanpa penyesuaian sintaks antar-basis data yang sering menimbulkan perbedaan hasil. Konsekuensi dari pilihan ini adalah artikel yang hanya terindeks pada basis data lain, misalnya Web of Science atau ACM Digital Library, berpotensi tidak tertangkap, dan keterbatasan tersebut dinyatakan kembali pada bagian keterbatasan.

Setelah rekaman diperoleh, kriteria inklusi dan eksklusi diterapkan untuk menyeleksi studi yang relevan dengan tujuan penelitian. Kriteria inklusi didefinisikan sebagai berikut: studi (I1) membahas sistem rekomendasi, rekomendasi yang dipersonalisasi, atau rekomendasi produk; (I2) menangani *user cold-start*, rekomendasi *cold-start*, *sparsity*, atau keterbatasan data interaksi pengguna; (I3) menggunakan atau membahas data tekstual seperti ulasan, deskripsi produk, metadata, atribut produk, teks semantik, atau kata kunci preferensi; (I4) relevan dengan *e-commerce*, belanja daring, ritel daring, atau rekomendasi produk; dan (I5) melaporkan metode, model, kerangka, atau hasil eksperimen yang dapat digunakan dalam sintesis. Kriteria eksklusi didefinisikan sebagai berikut: studi dikeluarkan jika (E1) hanya membahas sistem rekomendasi secara umum tanpa relevansi *cold-start*,

sparsity, atau personalisasi; (E2) hanya menggunakan data *rating*, data demografi, log sesi, citra produk, atau data *multimodal* tanpa bukti tekstual yang jelas; (E3) berada di luar konteks rekomendasi produk dan tidak memberikan wawasan yang dapat ditransfer untuk *user cold-start*; (E4) tidak menyediakan informasi metodologis atau eksperimental yang memadai; atau (E5) merupakan rekaman duplikat, laporan yang tidak dapat diakses, atau rekaman yang dikecualikan oleh aturan penyaringan. Studi dengan keselarasan parsial terhadap cakupan tinjauan diklasifikasikan sebagai bukti pendukung ketika berkontribusi pada pemahaman yang lebih luas mengenai mitigasi *cold-start*, metode rekomendasi *baseline*, atau strategi pemodelan yang dapat ditransfer.

Setelah tahap perencanaan, tinjauan dilaksanakan dengan cara mencari, menyaring, menyeleksi, dan mengklasifikasikan studi. Watase digunakan untuk mendukung proses penyaringan, klasifikasi otomatis, dan ekstraksi otomatis. Namun, keluaran otomatis tidak digunakan secara mekanis; data hasil ekstraksi diperiksa ulang agar sintesis akhir tetap selaras dengan tujuan penelitian. Pencarian awal menghasilkan 898 rekaman dari Scopus. Sebelum penyaringan, 195 rekaman duplikat dihapus, 252 rekaman dihapus oleh alat otomatisasi berdasarkan rentang tahun 2021-2026, 50 rekaman dihapus karena tidak memenuhi kriteria tingkat publikasi, dan 1 rekaman dihapus karena tidak menyediakan abstrak. Setelah proses ini, 400 rekaman tersisa untuk penyaringan. Dari rekaman tersebut, 308 dikeluarkan karena kurang sesuai dengan fokus penelitian. Sebanyak 92 laporan diupayakan untuk diperoleh, tetapi 36 laporan tidak dapat diperoleh, sehingga 56 laporan dinilai kelayakannya. Setelah penilaian kelayakan penuh, 2 laporan dikeluarkan, menghasilkan 55 laporan yang digunakan dalam ekstraksi dan sintesis akhir. Proses seleksi studi berdasarkan diagram alir PRISMA 2020 [56] disajikan pada Gambar 1.



Gambar 1 Proses seleksi studi berdasarkan diagram alir PRISMA 2020

Ke-55 artikel hasil ekstraksi kemudian diperiksa ulang berdasarkan relevansinya terhadap tujuan tinjauan dan dibagi menjadi studi inti (*core studies*) dan studi pendukung (*supporting studies*). Studi inti adalah studi yang memadukan data tekstual, relevansi *cold-start* atau *sparsity*, rekomendasi yang

dipersonalisasi, serta konteks *e-commerce* atau rekomendasi produk. Studi pendukung adalah studi yang membahas *cold-start*, metode sistem rekomendasi, rekomendasi *hybrid*, rekomendasi berbasis graf, *matrix factorization*, atau rekomendasi lintas domain, tetapi tidak secara langsung menggunakan data tekstual atau tidak berfokus pada rekomendasi produk *e-commerce*.

Penilaian kualitas studi dilakukan setelah tahap kelayakan untuk memastikan bahwa setiap studi yang masuk ke dalam sintesis menyediakan informasi minimum yang diperlukan untuk menjawab ketiga pertanyaan penelitian. Instrumen penilaian disusun mengikuti praktik penilaian kualitas pada panduan tinjauan sistematis [57] dan diturunkan menjadi lima kriteria yang dapat diverifikasi langsung pada berkas ekstraksi. Setiap kriteria dinilai secara biner, yaitu bernilai satu apabila informasi dinyatakan secara eksplisit dan bernilai nol apabila tidak, sehingga rentang skor tiap studi adalah nol sampai lima. Studi dipertahankan dalam sintesis apabila memenuhi empat kriteria inti, yaitu QA1, QA2, QA4, dan QA5. Kriteria QA3 tidak diperlakukan sebagai syarat gugur melainkan sebagai penanda kekuatan bukti, karena sebagian studi bersifat konseptual atau berupa tinjauan sehingga tidak melaporkan metrik eksperimen. Kriteria beserta hasil penerapannya disajikan pada Tabel 4.

Tabel 4 Kriteria penilaian kualitas studi dan hasil penerapannya

Kode	Kriteria penilaian	Studi memenuhi (n = 55)
QA1	Metode penelitian atau pendekatan pemodelan dinyatakan secara eksplisit.	55 (100%)
QA2	Sumber data yang digunakan dinyatakan secara eksplisit.	55 (100%)
QA3	Metrik evaluasi atau prosedur analisis kuantitatif dilaporkan.	43 (78,2%)
QA4	Hasil atau temuan utama dilaporkan secara eksplisit.	55 (100%)
QA5	Keluarga model dapat diidentifikasi dari uraian studi.	55 (100%)

Seluruh 55 studi memenuhi empat kriteria inti sehingga tidak ada studi yang gugur pada tahap ini. Kriteria QA3 dipenuhi oleh 43 studi atau 78,2%, sedangkan 12 studi lainnya memuat nama metode atau sintesis kualitatif pada bidang analisis; dari 12 studi tersebut, enam berperan sebagai studi inti dan enam sebagai studi pendukung. Karena sebagian ketidaklengkapan ini bersumber dari proses ekstraksi otomatis dan bukan dari artikel aslinya, metrik untuk seluruh model yang masuk ke dalam perbandingan pada bagian hasil ditelusuri ulang secara langsung ke artikel primer, sedangkan studi tanpa metrik eksplisit hanya digunakan untuk menjawab RQ1 dan RQ2 yang bersifat pemetaan dan tidak dijadikan dasar klaim performa pada RQ3.

Ekstraksi data dilakukan menggunakan berkas ekstraksi dan berkas klasifikasi Watase. Bidang ekstraksi meliputi Country, Context, Research Methods, Analysis, Software, Grand Theory, Novelty, Result, Limitation, dan Future Research. Berkas klasifikasi digunakan untuk mendukung pemetaan sumber data tekstual, kategori metode, teknik NLP, dan metrik evaluasi. Data hasil ekstraksi kemudian dinormalisasi sebelum sintesis. Sebagian hasil ekstraksi otomatis menempatkan nama metode pada bidang Analysis; dalam tinjauan ini bidang tersebut ditafsirkan secara cermat sehingga metrik evaluasi seperti RMSE, MAE, Precision, Recall, F1-score, NDCG, MAP, Hit Ratio, AUC, dan Coverage diperlakukan sebagai bukti performa, bukan nama model. Bidang ekstraksi data dirangkum pada Tabel 5.

Tabel 5 Formulir ekstraksi data

Item Ekstraksi	Deskripsi
Country	Negara atau konteks afiliasi utama studi.
Context	Domain aplikasi dan setting rekomendasi.
Research Methods	Metode AI, ML, DL, NLP, atau sistem rekomendasi yang digunakan.
Analysis	Metrik evaluasi atau analisis eksperimen yang dilaporkan studi.
Software	Perangkat lunak, kerangka, atau tool yang secara eksplisit disebutkan.
Grand Theory	Landasan teoretis atau metodologis utama.
Novelty	Kontribusi utama yang diklaim studi.

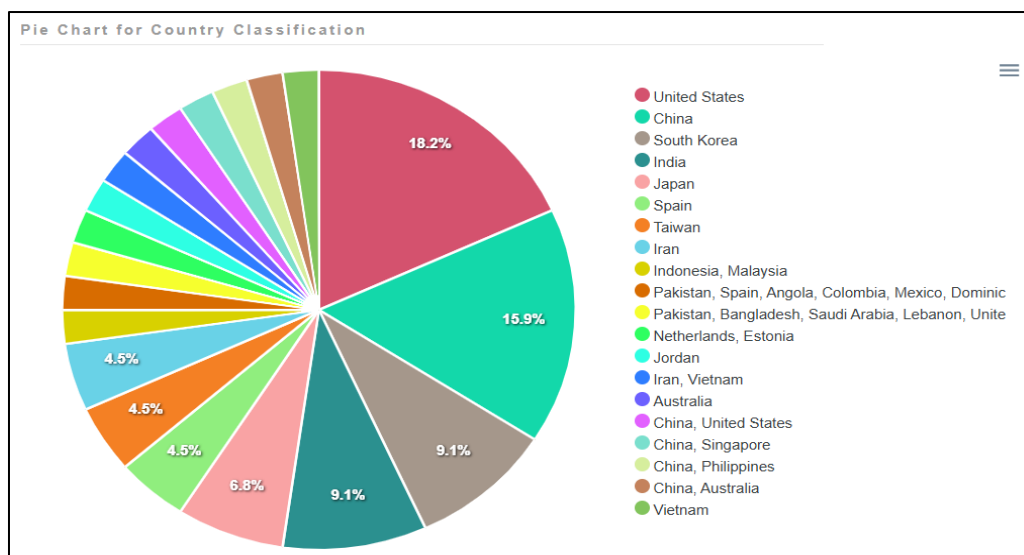
Item Ekstraksi	Deskripsi
Result	Temuan utama atau capaian performa.
Limitation	Keterbatasan yang dilaporkan studi.
Future Research	Arah penelitian lanjutan yang disarankan studi.

Sintesis dilakukan dalam tiga tahap. Pertama, sumber data tekstual diidentifikasi untuk menjawab RQ1, mencakup ulasan pengguna, ulasan item, deskripsi produk, judul produk, kategori produk, metadata produk, atribut produk, tag produk, query pencarian, kata kunci preferensi, dan sumber teks campuran. Kedua, metode AI, ML, DL, dan NLP dipetakan untuk menjawab RQ2, mencakup *content-based filtering*, *collaborative filtering*, rekomendasi *hybrid*, analisis sentimen, *topic modeling*, model berbasis *embedding*, model berbasis *transformer*, rekomendasi berbasis graf, dan rekomendasi berbasis ontologi. Ketiga, keluarga model dibandingkan untuk menjawab RQ3 dengan menekankan kesesuaian model terhadap *user cold-start*, bukan sekadar frekuensi metrik. Pada tahap pelaporan, studi inti digunakan sebagai bukti utama untuk menjawab pertanyaan penelitian, sedangkan studi pendukung digunakan untuk menyediakan latar belakang, pembandingan *baseline*, atau konteks sistem rekomendasi yang lebih luas.

4 Hasil dan Pembahasan

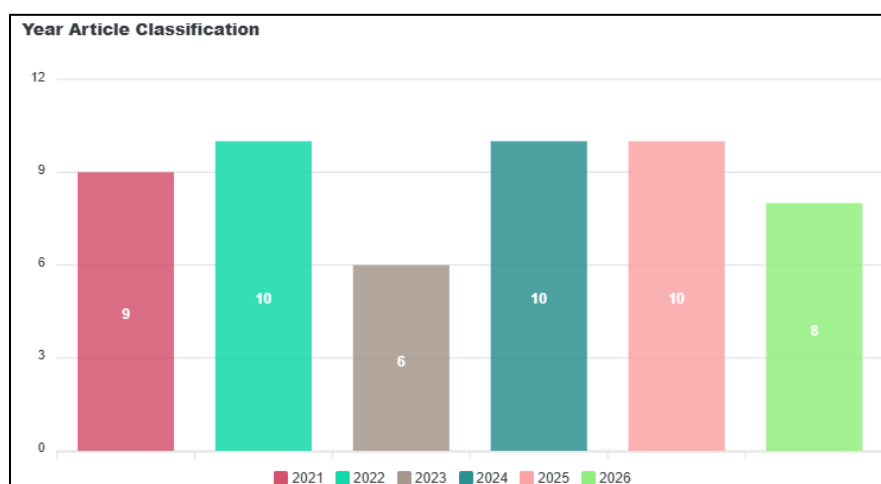
Korpus akhir tinjauan terdiri atas 55 artikel yang diperoleh dari proses ekstraksi dan klasifikasi Watase. Setelah pemeriksaan relevansi, studi terpilih dibagi menjadi studi inti dan studi pendukung. Studi inti memadukan data tekstual, relevansi *user cold-start* atau *sparsity*, rekomendasi yang dipersonalisasi, dan konteks *e-commerce* atau rekomendasi produk, sedangkan studi pendukung membahas *cold-start*, metode sistem rekomendasi, rekomendasi *hybrid*, rekomendasi berbasis graf, *matrix factorization*, atau rekomendasi lintas domain tanpa secara langsung menggunakan data tekstual.

Berdasarkan klasifikasi negara dari hasil ekstraksi Watase, studi terpilih tersebar di beberapa negara dan afiliasi multi-negara. Proporsi terbesar berasal dari Amerika Serikat sebesar 18,2% dari rekaman terklasifikasi, diikuti Tiongkok sebesar 15,9%, Korea Selatan dan India masing-masing 9,1%, Jepang 6,8%, serta Spanyol, Taiwan, dan Iran masing-masing 4,5%. Studi lainnya tersebar di kelompok negara yang lebih kecil dengan kontribusi sekitar 2,3% masing-masing, termasuk Indonesia dan Malaysia, Pakistan, Bangladesh, Arab Saudi, Lebanon, Uni Emirat Arab, Belanda dan Estonia, Yordania, Iran dan Vietnam, Australia, serta beberapa afiliasi gabungan yang melibatkan Tiongkok. Distribusi ini menunjukkan bahwa riset sistem rekomendasi dan *cold-start* dilakukan pada konteks geografis yang beragam, meskipun kontribusi terbesar masih terpusat di Amerika Serikat dan Tiongkok. Distribusi negara studi terpilih disajikan pada Gambar 2.



Gambar 2 Klasifikasi negara studi terpilih

Klasifikasi tahun menunjukkan bahwa studi terpilih tersebar pada publikasi tahun 2021 hingga 2026. Jumlah artikel tertinggi muncul pada 2022, 2024, dan 2025, masing-masing 10 artikel. Tahun 2021 menyumbang 9 artikel, diikuti 2026 dengan 8 artikel, dan 2023 dengan 6 artikel. Distribusi ini menunjukkan bahwa riset mengenai rekomendasi *cold-start* dan sistem rekomendasi berbasis teks tetap aktif dalam beberapa tahun terakhir, dengan konsentrasi yang relatif kuat antara 2022 dan 2025. Distribusi tahun publikasi disajikan pada Gambar 3, sedangkan ikhtisar bukti yang diekstraksi dirangkum pada Tabel 6.



Gambar 3 Klasifikasi tahun artikel terpilih berdasarkan ekstraksi watase

Tabel 6 Ikhtisar studi terpilih dan bukti yang diekstraksi

Aspek	Hasil
Artikel akhir yang diekstraksi	55 artikel.
Studi inti	31 studi.
Studi pendukung	24 studi.
Klasifikasi negara	Amerika Serikat (18,2%), Tiongkok (15,9%), Korea Selatan (9,1%), India (9,1%), Jepang (6,8%), Spanyol (4,5%), Taiwan (4,5%), Iran (4,5%), dan afiliasi tunggal atau multi-negara lain (2,3% masing-

Aspek	Hasil
	masing).
Klasifikasi tahun	2021 (9 artikel), 2022 (10 artikel), 2023 (6 artikel), 2024 (10 artikel), 2025 (10 artikel), dan 2026 (8 artikel).

Sumber data tekstual pada user cold-start (RQ1). Studi terpilih menunjukkan bahwa data tekstual penting untuk mengurangi keterbatasan riwayat interaksi yang jarang pada rekomendasi *user cold-start*. Ketika pengguna baru belum menghasilkan cukup *rating*, klik, pembelian, atau pola penelusuran, informasi tekstual dapat menyediakan sinyal lebih awal untuk merepresentasikan produk dan menyimpulkan preferensi pengguna. Ulasan pengguna merupakan sumber tekstual yang paling sering digunakan pada bukti inti karena memuat opini, pengalaman produk, sentimen, dan penilaian tingkat aspek yang tidak sepenuhnya terwakili oleh *rating* numerik; DeepSentRec, misalnya, menggunakan informasi sentimen berbasis ulasan untuk mendukung rekomendasi produk pada kondisi *cold-start* dan data jarang [8].

Ulasan item juga memberikan bukti yang berguna karena merangkum bagaimana produk dipersepsikan oleh banyak pengguna, dan SuReFAR menunjukkan bagaimana ringkasan ulasan dapat memperkuat rekomendasi dengan mengurangi derau dari teks ulasan yang panjang [14]. Deskripsi produk merupakan sumber tekstual penting lain karena tersedia sebelum pengguna berinteraksi dengan produk; deskripsi dapat disandikan menjadi representasi semantik dan dibandingkan meskipun riwayat interaksi terbatas, sebagaimana representasi deskripsi item berbasis SBERT [9]. Metadata produk seperti kategori, atribut, merek, dan tag turut berkontribusi, dan OntoCommerce menunjukkan bahwa atribut produk dan relasi semantik dapat dipadukan ke dalam model rekomendasi *e-commerce* berbasis ontologi [21]. Kata kunci preferensi dan teks kontekstual juga dapat mendukung pemodelan awal pengguna, dan Context2Rec menunjukkan bahwa semantik teks kontekstual dapat digunakan bersama fitur sekuensial untuk meningkatkan rekomendasi [16]. Secara keseluruhan, sumber tekstual yang paling menonjol pada studi inti adalah ulasan pengguna, ulasan item, deskripsi produk, metadata dan atribut produk, serta teks terkait preferensi, sementara data *rating*, demografi, log sesi, citra produk, dan data *multimodal* muncul terutama sebagai bukti pendukung. Beberapa studi juga menggunakan bentuk tekstual yang lebih khusus, seperti analisis sentimen semantik, analisis emosi tingkat kalimat, dan analisis sentimen *hybrid* [39], [45], [50], [52], atau memadukan teks dengan relasi kepercayaan, relasi item semantik, dan *query* bahasa alami untuk rekomendasi yang *trust-aware*, semantik, dan dapat dijelaskan [49], [51].

Metode AI, ML, DL, dan NLP (RQ2). Studi terpilih menunjukkan bahwa masalah *user cold-start* umumnya ditangani dengan metode rekomendasi *hybrid*, karena skenario pengguna baru jarang menyediakan satu sinyal yang kuat. Oleh karena itu, informasi tekstual sering dikombinasikan dengan *collaborative filtering*, *matrix factorization*, *neural representation learning*, analisis sentimen, pemodelan graf, atau penalaran berbasis ontologi. Rekomendasi *sentiment-aware* merupakan salah satu keluarga metode paling relevan untuk *cold-start* berbasis teks karena mengekstraksi opini atau polaritas emosi dari ulasan; DeepSentRec adalah contoh model *deep learning sentiment-aware* untuk meningkatkan rekomendasi produk menggunakan sentimen berbasis ulasan [8].

Representasi berbasis *embedding* juga sering digunakan karena data tekstual harus dikonversi menjadi vektor numerik sebelum dapat digunakan model rekomendasi; model berbasis SBERT menyandikan deskripsi item menjadi representasi semantik yang dapat dibandingkan antar-produk, dan Lim serta Lee menunjukkan bahwa *embedding* SBERT dan fitur sentimen dapat memperkuat rekomendasi berbasis MF dan NCF [9]. Model *transformer* dan *attention* penting karena mampu menangkap makna kontekstual dari teks ulasan dan deskripsi produk; SuReFAR menggunakan peringkasan ulasan dan *neural representation learning* untuk meningkatkan rekomendasi dari bukti tekstual [14], sedangkan SRARS menggunakan peringkasan ulasan dan mekanisme *attention* untuk mendukung rekomendasi dari data ulasan [37]. Rekomendasi berbasis graf berguna ketika informasi tekstual dapat dihubungkan dengan relasi pengguna-produk, atribut produk, atau relasi pengetahuan; Talk-with-Graph memadukan pemahaman teks berbasis LLM, analisis sentimen, *fuzzy logic*, dan heterogeneous *graph neural network* [38], sementara RAKCR menggunakan sentimen ulasan dan *knowledge graph convolutional network* untuk memperkaya representasi produk dan pengguna [18]. Metode AI tambahan muncul sebagai arah pendukung, seperti model *cross-domain* yang mentransfer

preferensi tingkat aspek [30], [35], [42], serta model berbasis *hypernetwork*, GAN, *attention*, dan multi-level latent factor untuk kondisi *cold-start* atau data jarang [43], [46], [48], [5]. Rekomendasi berbasis ontologi di luar konteks *e-commerce* turut memberikan bukti pendukung bahwa pengetahuan domain dapat mengurangi masalah pengguna baru [12], sedangkan model *multimodal* dan *visual-semantic* memperluas gagasan tersebut dengan menggabungkan teks dan citra, meskipun diperlakukan sebagai bukti pendukung karena fokus tinjauan ini tetap pada rekomendasi berbasis teks [44], [54]. Skema *hybrid cross-breed* pada *e-commerce* juga dilaporkan menangani *cold-start* dan *sparsity* melalui penggabungan beberapa sumber sinyal [55]. Pola metode yang berulang pada studi terpilih mengarah pada kombinasi *hybrid*, *sentiment-aware*, representasi *embedding*, NLP berbasis *transformer* atau *attention*, peringkasan ulasan, serta rekomendasi berbasis graf dan ontologi; *collaborative filtering* dan *matrix factorization* murni tetap dipakai sebagai *baseline*, namun jarang berdiri sendiri sebagai solusi *cold-start*.

Kandidat model dengan performa terbaik yang dilaporkan (RQ3). Sintesis RQ3 memadukan kesesuaian model dengan bukti empiris yang dilaporkan masing-masing studi, karena efektivitas pada rekomendasi *user cold-start* bergantung pada capaian performa sekaligus keselarasan dengan kondisi data pengguna yang jarang. Metrik evaluasi seperti RMSE, MAE, Precision, Recall, NDCG, Hit Rate, AUC, dan F1-score bukanlah model itu sendiri, tetapi penting untuk mendukung klaim efektivitas. Oleh karena itu, bagian ini mengidentifikasi model efektif dengan mempertimbangkan data tekstual yang digunakan, jenis masalah *cold-start* atau *sparsity* yang ditangani, *dataset*, metrik evaluasi, hasil yang dilaporkan, dan relevansi tiap model terhadap rekomendasi produk *e-commerce*. Perbandingan model efektif disajikan pada Tabel 7.

Tabel 7 Perbandingan model efektif untuk rekomendasi produk *user cold-start*

Model	Data Tekstual	Dataset	Metrik	Hasil yang Dilaporkan	Relevansi
DeepSentRec [8]	Ulasan pengguna dan ulasan item dengan informasi sentimen	Amazon Reviews, Yelp, IMDB, Sentiment, dan Kaggle E-Commerce Reviews	Precision@10, Recall@10, NDCG@10, MAP, Hit Rate, CTR	Dilaporkan mengungguli CF, SBERT-only, dan HybridCF-SBERT. Pada Amazon: Precision@10 = 91,12%, Recall@10 = 85,49%, NDCG@10 = 89,66%, CTR = 24,96%. Pada Kaggle: Precision@10 = 89,84%, Recall@10 = 83,92%, NDCG@10 = 87,90%, CTR = 23,41%.	Sangat relevan karena langsung menggunakan sentimen ulasan untuk rekomendasi produk pada kondisi <i>cold-start</i> dan data jarang.
SBERT+VADER-enhanced MF/NCF [9]	Deskripsi item disandikan SBERT dan teks ulasan dianalisis VADER	Amazon Grocery and Gourmet Food, Amazon Video Games	Recall@10, NDCG@10, HIT@10	Model side-information memperbaiki mayoritas <i>baseline</i> . Untuk Video Games, MF meningkat RMSE 3,111 -> 1,163; Recall@10 0,482 -> 0,501; NDCG@10 0,285 -> 0,296; HIT@10 0,479 -> 0,498. NCF meningkat RMSE 1,151 -> 1,130; Recall@10 0,448 -> 0,479; NDCG@10 0,254 -> 0,300; HIT@10 0,446 -> 0,477.	Sangat relevan karena memadukan makna objektif deskripsi dan emosi subjektif ulasan.
SuReFAR [14]	Teks ulasan dan ringkasan ulasan	Amazon Books, Amazon Movie and TV, Yelp	MAE, RMSE, MSE, paired t-test	Dilaporkan mengungguli NCF, DeepCoNN, NARRE, AENAR, dan MFNR. MAE/RMSE = 0,476/0,757 (Amazon Books), 0,696/1,070 (Amazon Movie and TV), 0,820/1,068 (Yelp).	Sangat relevan karena memakai peringkasan ulasan untuk mengurangi derau dan memperkuat rekomendasi <i>review-aware</i> .
SRARS [37]	Teks ulasan dan ulasan yang diringkas	Amazon Video Games, Amazon Industrial and Scientific, Amazon Musical Instruments	MAE, RMSE	Dilaporkan mengungguli NCF, D-attn, DeepCoNN, NARRE, dan AENAR. MAE membaik 3,44-23,66% (Video Games), 5,91-25,00% (Industrial and Scientific), 6,54-21,32% (Musical Instruments).	Relevan karena memadukan peringkasan dan <i>attention</i> untuk mengekstraksi sinyal preferensi yang lebih kuat.
Context2Rec [16]	Semantik teks komentar dan teks kontekstual	Amazon, Yelp, dan Beer	MSE, MAE	Dilaporkan mengungguli PMF, LFM, SVD++, NARRE, DeepCoNN, SASRec, dan BERT4Rec. MSE berkurang 14,3% (Amazon), 16,2% (Yelp), 12,6% (Beer).	Relevan karena memadukan semantik teks dan fitur sekuensial untuk rekomendasi yang dipersonalisasi.
Talk-with-Graph [38]	Teks ulasan, sentimen berbasis LLM,	Amazon Musical Instruments	MRR@10, NDCG@10, HIT@10,	Dilaporkan mengungguli metode <i>baseline</i> pada berbagai metrik. Pada Amazon Musical Instruments: MRR@10	Relevan ketika relasi produk-pengguna dapat

Model	Data Tekstual	Dataset	Metrik	Hasil yang Dilaporkan	Relevansi
	dan relasi graf		Precision@10, Recall@10, F1-score	= 0,0681, NDCG@10 = 0,0984, HIT@10 = 0,0195, Precision@10 = 0,0195, Recall@10 = 0,0997. Modul sentimen: akurasi 0,973, presisi 0,972, recall 0,974, F1 0,973.	direpresentasikan sebagai graf dan dipadukan dengan penalaran sentimen berbasis teks.
OntoCommerce [21]	Deskripsi produk, atribut produk, dan relasi semantik produk	Dataset rekomendasi produk <i>e-commerce</i> yang digunakan pada studi	MAE, Precision, Recall, kepuasan pelanggan	Dilaporkan mengungguli CF dan CF+ <i>Ontology</i> . Pada ukuran tetangga 20: MAE = 0,66 (CF+ <i>Ontology</i> 0,69; CF 0,76). Kepuasan pelanggan sekitar 95% (CF+ <i>Ontology</i> 78%; CF 60%).	Sangat relevan ketika pengetahuan produk, atribut, dan relasi semantik tersedia.
RAKCR [18]	Sentimen ulasan dan informasi <i>knowledge graph</i>	MovieLens-20M, Amazon-Book, Yelp	AUC, F1-score, Recall@K	Dilaporkan mengungguli BPRMF, CFKG, RippleNet, KGCN, KGAT, LightGCN, dan KGCL. Untuk prediksi CTR: AUC/F1 = 0,943/0,873 (MovieLens-20M), 0,946/0,896 (Amazon-Book), 0,947/0,941 (Yelp).	Relevan sebagai bukti pendukung karena memadukan sentimen ulasan dan <i>knowledge graph convolution</i> untuk memperkaya representasi.

Perbandingan pada Tabel 7 menunjukkan bahwa model paling efektif tidak ditentukan oleh satu metrik evaluasi saja, melainkan oleh kemampuan memadukan bukti tekstual dengan sinyal rekomendasi lain. DeepSentRec dan SBERT+VADER-enhanced MF/NCF menonjol karena memadukan sentimen ulasan dan representasi semantik produk dengan rekomendasi *hybrid*. SuReFAR dan SRARS menjadi kandidat kuat ketika sumber tekstual utama adalah teks ulasan, terutama pada ulasan yang panjang atau bising. Context2Rec berguna ketika konteks tekstual dan perilaku sekuensial tersedia, sedangkan Talk-with-Graph, OntoCommerce, dan RAKCR lebih sesuai ketika informasi tekstual dapat dihubungkan dengan relasi graf, ontologi, atau pengetahuan. Berdasarkan perbandingan ini, arah yang paling dapat dipertahankan untuk *user cold-start* pada rekomendasi produk *e-commerce* adalah sistem rekomendasi *hybrid text-semantic*. Perlu ditegaskan bahwa tinjauan ini tidak melakukan eksperimen pembandingan langsung; namun pada studi-studi terpilih, pendekatan semacam ini cenderung dilaporkan lebih unggul dibanding *collaborative filtering* atau *matrix factorization* murni karena dapat memanfaatkan informasi tekstual sebelum riwayat interaksi memadai. Namun capaian yang dilaporkan tersebut baru dapat dibaca secara utuh apabila disandingkan dengan keterbatasan yang diakui oleh studi asal dan dengan kondisi data yang menjadi prasyarat penerapannya, sebagaimana dirangkum pada Tabel 8.

Tabel 8 Analisis kritis antar-keluarga model: kelebihan, keterbatasan terlapor, dan kondisi penerapan

Keluarga model	Kelebihan utama	Keterbatasan yang dilaporkan studi asal	Kondisi penerapan
Rekomendasi <i>hybrid deep sentiment-aware</i>	Menangkap opini, emosi, dan penilaian tingkat aspek yang tidak terwakili oleh <i>rating</i> numerik, sehingga profil pengguna baru dapat dibentuk segera setelah ulasan tersedia.	DeepSentRec bergantung pada label sentimen tersupervisi yang belum tentu tersedia dan konsisten antar-domain, serta diuji terbatas pada ulasan berbahasa Inggris [8]. Model SBERT+VADER menggunakan model umum tanpa penyetelan domain dan masih menyisakan kesulitan pada kondisi <i>sparsity</i> dan <i>cold-start</i> [9].	Platform dengan volume ulasan besar, bahasa relatif seragam, serta kapasitas pelabelan atau penyetelan model pada domain sendiri.
Rekomendasi <i>transformer</i> atau <i>attention</i> berbasis ulasan	Meringkas ulasan yang panjang dan berulang menjadi representasi padat sehingga derau ditekan sebelum prediksi	SuReFAR membentuk profil pengguna dan item yang statis tanpa memperhitungkan dinamika temporal dan hanya mengolah teks ulasan [14]. SRARS hanya menggunakan teks ulasan tanpa	Katalog dengan ulasan panjang atau redundan, kebutuhan prediksi <i>rating</i> , dan ketersediaan sumber daya komputasi untuk

<http://sistemasi.ftik.unisi.ac.id>

Keluarga model	Kelebihan utama	Keterbatasan yang dilaporkan studi asal	Kondisi penerapan
	dilakukan.	profil pengguna atau deskripsi item dan divalidasi terbatas pada <i>dataset</i> Amazon [37]. Context2Rec bergantung pada ketersediaan ulasan dan diuji pada tiga <i>dataset</i> berbasis ulasan sehingga generalisasinya masih perlu pengujian lanjutan [16].	model <i>transformer</i> .
Rekomendasi semantik berbasis graf atau pengetahuan	Menautkan teks dengan relasi produk, atribut, dan pengetahuan sehingga penalaran rekomendasi tetap dimungkinkan ketika interaksi pengguna sangat minim.	Talk-with-Graph sangat bergantung pada rancangan <i>prompt</i> , menuntut pembaruan bobot dinamis serta pemeliharaan graf berskala besar, dan berbiaya komputasi tinggi [38]. OntoCommerce diuji pada <i>dataset</i> yang relatif kecil dan bertumpu pada ontologi, <i>rating</i> , serta pola pembelian, bukan pada deskripsi produk tekstual [21]. RAKCR membatasi fitur sentimen pada modalitas teks sehingga generalisasi lintas-domain belum terjamin [18].	Katalog terstruktur dengan ontologi atau <i>knowledge graph</i> yang terpelihara serta kapasitas rekayasa data yang memadai.
Rekomendasi <i>hybrid</i> berbasis <i>embedding</i> atau topik	Lebih ringan secara komputasi dan dapat berperan sebagai <i>baseline</i> yang kuat ketika sumber daya untuk model <i>transformer</i> terbatas.	Pendekatan <i>topic embedding</i> bergantung pada ketersediaan dan kualitas ulasan sehingga performanya menurun ketika teks terbatas, sangat pendek, atau bising [34]. Model berbasis topik juga bergantung pada informasi tambahan seperti karakteristik pengguna dan fitur item, dan varian LDA-HDP sulit diterapkan karena menghasilkan dimensi topik yang besar [33]. Model <i>word sequential LSTM</i> +MF belum menguji pengaruh <i>hyperparameter</i> maupun skalabilitas pada <i>dataset</i> yang lebih besar [10].	Platform dengan sumber daya komputasi terbatas dan kebutuhan <i>baseline</i> yang mudah diinterpretasikan.
Baseline CF/MF/NCF murni	Sederhana, terstandar, dan berguna sebagai pembanding yang jelas dalam pengukuran performa.	Model RBPR dilaporkan belum bekerja baik pada umpan balik eksplisit yang sangat jarang meskipun telah dipraproses dengan SVD [6], dan pendekatan jenis ini tetap memerlukan bukti interaksi atau <i>rating</i> sehingga kurang sesuai sebagai solusi tunggal bagi pengguna baru [5].	Digunakan sebagai pembanding, atau ketika data interaksi pengguna telah cukup padat.

Pembacaan Tabel 8 memperlihatkan bahwa keunggulan tiap keluarga model selalu disertai prasyarat data yang berbeda, sehingga pemilihan model tidak dapat dilepaskan dari kondisi platform. Keluarga *sentiment-aware* menuntut ulasan yang melimpah dan relatif seragam secara bahasa; keluarga *review-aware* menuntut teks yang cukup panjang untuk diringkas; keluarga berbasis graf dan

ontologi menuntut struktur relasi yang terpelihara beserta biaya komputasi yang tidak kecil; sedangkan keluarga berbasis *embedding* dan topik menuntut sumber daya paling sedikit tetapi paling sensitif terhadap ulasan yang pendek atau bising. Pola ini menjelaskan mengapa tidak ada satu keluarga model yang unggul pada seluruh kondisi, dan mengapa keterbatasan yang diakui studi asal justru menjadi penentu kelayakan penerapan pada praktik.

Perlu ditegaskan pula bahwa angka pada Tabel 7 tidak dapat diperlakukan sebagai peringkat performa antar model. Studi-studi tersebut dievaluasi pada *dataset* yang berbeda, misalnya Amazon Books, Amazon Video Games, Yelp, MovieLens-20M, dan *dataset e-commerce* internal, dengan protokol pembagian data serta definisi skenario *cold-start* yang tidak seragam. Metrik yang dilaporkan pun berbeda jenis dan tidak sebanding secara langsung: nilai kesalahan prediksi seperti MAE 0,476 pada SuReFAR tidak berada pada skala yang sama dengan AUC 0,943 pada RAKCR maupun Precision@10 sebesar 91,12% pada DeepSentRec, karena ketiganya mengukur hal yang berbeda pada tugas dan *dataset* yang berbeda. Metrik bernama sama juga dapat dihitung pada nilai potong K yang berbeda, dan setiap studi membandingkan dirinya terhadap himpunan *baseline* yang dipilih sendiri. Oleh karena itu, angka pada Tabel 7 dibaca sebagai bukti internal setiap studi terhadap *baseline*-nya sendiri, bukan sebagai bukti lintas-studi, dan sintesis RQ3 pada tinjauan ini bertumpu pada kesesuaian model terhadap kondisi data, bukan pada peringkat metrik. Ringkasan keluarga model beserta kondisi yang paling sesuai disajikan pada Tabel 9.

Tabel 9 Rekomendasi keluarga model untuk rekomendasi produk user cold-start

Keluarga Model	Model Representatif	Interpretasi
Rekomendasi <i>hybrid deep sentiment-aware</i>	DeepSentRec, SBERT+VADER-enhanced MF/NCF	Paling sesuai ketika ulasan pengguna atau ulasan item tersedia karena model dapat menangkap opini, emosi, dan petunjuk preferensi.
Rekomendasi <i>transformer</i> atau <i>attention</i> berbasis ulasan	SuReFAR, SRARS, Context2Rec	Paling sesuai ketika ulasan, deskripsi, atau teks kontekstual dapat diringkas dan disandikan sebagai representasi semantik.
Rekomendasi semantik berbasis graf atau pengetahuan	Talk-with-Graph, OntoCommerce, RAKCR	Paling sesuai ketika relasi produk, atribut, ontologi, atau <i>knowledge graph</i> tersedia.
Rekomendasi <i>hybrid</i> berbasis <i>embedding</i> atau topik	<i>Topic embedding</i> model, word sequential LSTM+MF	Berguna sebagai keluarga model yang lebih ringan atau <i>baseline</i> ketika sumber daya <i>transformer</i> terbatas.
<i>Baseline</i> CF/MF/NCF murni	<i>Matrix factorization</i> , <i>collaborative filtering</i> , NCF, RBPR	Berguna untuk pembandingan, tetapi kurang sesuai sebagai satu-satunya solusi untuk <i>user cold-start</i> berbasis teks.

Temuan tinjauan ini menunjukkan bahwa data tekstual berperan penting dalam mengurangi kelemahan sistem rekomendasi pada kondisi *user cold-start*. Pada skenario pengguna baru, sistem tidak dapat mengandalkan riwayat interaksi jangka panjang, sehingga informasi tekstual menjadi sumber alternatif untuk merepresentasikan produk dan menyimpulkan preferensi awal. Hal ini menjelaskan mengapa ulasan, deskripsi produk, metadata, atribut produk, dan informasi semantik produk muncul sebagai sumber penting. Hasil juga menunjukkan bahwa rekomendasi *cold-start* berbasis teks tidak seharusnya dipahami sebagai masalah metode tunggal; model yang hanya menggunakan *collaborative filtering* atau *matrix factorization* melemah ketika riwayat pengguna jarang, sehingga arah yang lebih meyakinkan adalah pendekatan *hybrid text-semantic* yang memadukan representasi tekstual dengan fitur sentimen, *neural* ranking, relasi graf, atau pengetahuan berbasis ontologi. Efektivitas model bergantung pada ketersediaan data: ketika data ulasan kaya, model *sentiment-aware* dan *review-aware* lebih sesuai; ketika deskripsi produk lebih lengkap, representasi semantik berbasis BERT atau SBERT lebih praktis; dan ketika katalog produk memuat relasi yang jelas, rekomendasi berbasis graf atau ontologi memberikan struktur semantik yang lebih kuat. Interpretasi ini turut didukung oleh model yang memadukan sentimen ulasan dengan *rating*, sentimen semantik dengan perilaku transaksi, atau kepuasan pelanggan dengan data penjualan [39], [40], [45], [47].

Temuan tinjauan ini konsisten dengan studi sebelumnya yang mengidentifikasi *cold-start* sebagai masalah yang persisten. Yuan dan Hernandez [4] menjelaskan bahwa *user cold-start* tetap sulit karena pengguna baru memberikan informasi terbatas; tinjauan ini mendukung pandangan tersebut namun menambahkan interpretasi yang lebih spesifik, yaitu data tekstual dapat menjadi jembatan praktis antara pengguna baru dan representasi produk pada *e-commerce*. Ray [19] meninjau pendekatan dan algoritma mitigasi *cold-start* secara lebih luas; tinjauan ini mempersempit pembahasan dengan menekankan bukti berbasis teks dan menunjukkan bahwa model *hybrid* lebih kuat ketika mengintegrasikan teks produk, sentimen ulasan, dan representasi semantik. Vahidy dkk. [5] membahas sistem rekomendasi dan *cold-start* secara umum, membantu menjelaskan keterbatasan sistem konvensional, sementara tinjauan ini berfokus lebih langsung pada bagaimana sumber tekstual dapat mengompensasi data interaksi yang jarang. Raeesi dkk. [20] menunjukkan bahwa *text mining* dapat mengidentifikasi tren riset, tetapi tinjauan ini berbeda karena mengevaluasi kontribusi data tekstual dan model berbasis NLP terhadap penyelesaian *user cold-start*, bukan sekadar identifikasi tren. Kontribusi tinjauan ini terletak pada fokus yang lebih sempit dan aplikatif, yakni memetakan sumber data tekstual, metode AI/NLP, dan keluarga model yang relevan untuk *user cold-start* pada rekomendasi produk *e-commerce*.

Pemilihan model sebaiknya dimulai dari jenis data tekstual yang tersedia pada platform *e-commerce*. Jika platform memiliki banyak ulasan pengguna atau ulasan item, model *hybrid deep sentiment-aware* sesuai karena dapat mengekstraksi opini, emosi, dan sinyal preferensi tingkat aspek, sebagaimana DeepSentRec [8]. Jika platform memiliki deskripsi produk yang kaya tetapi data ulasan terbatas, model *embedding* semantik lebih sesuai; representasi berbasis SBERT dapat mengubah deskripsi item menjadi vektor semantik yang dapat dibandingkan meskipun pengguna baru belum menghasilkan interaksi memadai [9]. Jika platform memiliki teks ulasan yang panjang dan bising, model *transformer* atau *attention* berbasis ulasan lebih sesuai, seperti SuReFAR dan SRARS yang menunjukkan bahwa peringkasan ulasan dan *attention* dapat mengubah teks ulasan panjang menjadi sinyal rekomendasi yang lebih berguna [14], [37]. Jika relasi produk, struktur kategori, atau koneksi atribut tersedia, model berbasis graf dan ontologi memberikan nilai tambah, sebagaimana Talk-with-Graph [38] dan OntoCommerce [21]. Dengan demikian, tidak ada satu model terbaik untuk setiap kondisi; rekomendasi paling dapat dipertahankan adalah arsitektur *hybrid* yang adaptif terhadap ketersediaan data.

Tinjauan ini memiliki beberapa keterbatasan. Pertama, studi terpilih diperoleh dari alur ekstraksi dan klasifikasi Watase, sehingga kualitas *dataset* awal bergantung pada strategi pencarian, kualitas metadata, dan konfigurasi penyaringan; meskipun studi diperiksa ulang secara manual, beberapa artikel relevan di luar cakupan pencarian mungkin tidak tercakup. Kedua, klasifikasi studi inti dan pendukung melibatkan penilaian interpretatif, dan sebagian studi tidak secara eksplisit memisahkan *user cold-start* dari *item cold-start* atau *sparsity* umum, sehingga menimbulkan ancaman terhadap validitas konstruk. Ketiga, studi terpilih menggunakan *dataset*, protokol evaluasi, dan metrik yang berbeda sehingga perbandingan model tidak dapat dilakukan sebagai perbandingan statistik langsung; efektivitas ditafsirkan berdasarkan kesesuaian metodologis, hasil eksperimen yang dilaporkan, dan relevansi terhadap *user cold-start*. Keempat, sebagian bidang ekstraksi otomatis tidak lengkap atau tidak konsisten, misalnya bidang Analysis kadang memuat nama metode alih-alih metrik evaluasi, sehingga informasi ditafsirkan secara cermat dan metrik hanya diperlakukan sebagai bukti pendukung. Kelima, beberapa studi pendukung berasal dari domain non-*e-commerce* atau menggunakan data non-tekstual seperti *rating*, log sesi, data demografi, atau citra produk, yang berfungsi sebagai bukti kontekstual untuk memosisikan rekomendasi berbasis teks dalam lanskap yang lebih luas.

Penelitian lanjutan sebaiknya mengembangkan *dataset benchmark* yang secara eksplisit memisahkan *cold-start* pengguna baru dari skenario item baru dan *sparsity* umum, karena model yang baik pada *sparsity* umum belum tentu menyelesaikan masalah pengguna baru. Studi berikutnya juga sebaiknya memadukan beberapa sumber tekstual dalam satu desain eksperimen, sehingga kontribusi deskripsi produk, metadata, ulasan pengguna, ulasan item, dan kata kunci preferensi dapat diamati lebih jelas. Peneliti perlu melaporkan *ablation study* secara lebih konsisten untuk menentukan apakah peningkatan performa berasal dari representasi tekstual, fitur sentimen, encoding berbasis *transformer*, struktur graf, pengetahuan ontologi, atau kombinasi *hybrid* itu sendiri. Penelitian lanjutan juga sebaiknya membandingkan model berbasis teks terhadap *baseline* yang jelas seperti TF-

<http://sistemasi.ftik.unisi.ac.id>

IDF, Word2Vec, LDA, *matrix factorization*, *neural collaborative filtering*, dan rekomendasi berbasis graf pada split *cold-start* yang sama, serta mengkaji apakah model *transformer* berbasis RAG, model sekuensial bergaya BERT, transfer preferensi tingkat aspek, dan Bayesian sentiment ranking tetap efektif pada split pengguna baru yang sama [29], [41], [42], [53]. Untuk implementasi praktis, sistem masa depan sebaiknya dirancang sebagai recommender *hybrid* adaptif yang memanfaatkan deskripsi dan metadata ketika data ulasan terbatas, model *review-aware* ketika data ulasan kaya, dan penalaran berbasis graf atau ontologi ketika relasi produk terstruktur tersedia.

5 Kesimpulan

Penelitian ini melakukan tinjauan literatur sistematis terhadap sistem rekomendasi berbasis teks untuk menangani masalah *user cold-start* pada rekomendasi produk *e-commerce*. Berdasarkan 55 artikel terpilih, tinjauan menemukan bahwa data tekstual penting untuk mengurangi keterbatasan riwayat interaksi pengguna yang jarang. Ulasan pengguna, ulasan item, deskripsi produk, metadata produk, atribut produk, dan informasi semantik produk merupakan sumber tekstual paling relevan karena dapat merepresentasikan produk dan menyimpulkan preferensi pengguna sebelum data perilaku memadai tersedia.

Tinjauan ini juga menunjukkan bahwa *user cold-start* pada rekomendasi *e-commerce* lebih baik ditangani melalui pendekatan *hybrid text-semantic* dibandingkan model rekomendasi metode tunggal. Model *hybrid deep sentiment-aware*, model *transformer* atau *attention* berbasis ulasan, model *embedding* semantik, dan model berbasis graf atau ontologi merupakan arah yang paling menonjol pada literatur yang ditinjau. Model seperti DeepSentRec, SBERT+VADER-enhanced MF/NCF, SuReFAR, SRARS, Context2Rec, Talk-with-Graph, OntoCommerce, dan RAKCR menjadi kandidat relevan karena memadukan bukti tekstual dengan sinyal rekomendasi tambahan. Kontribusi utama penelitian ini adalah pemetaan terfokus terhadap sumber data tekstual, metode AI/NLP, dan keluarga model yang efektif untuk *user cold-start* pada rekomendasi produk *e-commerce* yang dipersonalisasi, dengan catatan bahwa model yang paling sesuai bergantung pada ketersediaan data: platform kaya ulasan dapat memprioritaskan model *sentiment-aware* dan *review-aware*, platform kaya deskripsi dapat memprioritaskan representasi semantik berbasis BERT atau SBERT, dan platform dengan relasi produk terstruktur dapat memprioritaskan rekomendasi berbasis graf atau ontologi.

Secara praktis, temuan ini dapat diterjemahkan menjadi urutan keputusan yang sederhana bagi pengembang platform *e-commerce*. Langkah pertama adalah menginventarisasi aset tekstual yang benar-benar dimiliki, yaitu volume dan panjang ulasan, kelengkapan deskripsi produk, serta kerapian metadata dan relasi kategori. Langkah kedua adalah memilih keluarga model yang prasyarat datanya terpenuhi, yakni model *sentiment-aware* ketika ulasan melimpah, model *review-aware* ketika ulasan panjang dan bising, model *embedding* semantik ketika deskripsi produk lebih lengkap daripada ulasan, serta model berbasis graf atau ontologi ketika relasi produk telah terstruktur. Langkah ketiga adalah menyiapkan mekanisme cadangan berbasis deskripsi dan metadata bagi pengguna yang belum memiliki jejak interaksi sama sekali, karena seluruh keluarga model tetap memerlukan sinyal awal. Urutan ini menempatkan ketersediaan data, dan bukan kebaruan arsitektur, sebagai penentu utama pilihan model.

Penelitian ini memiliki keterbatasan karena artikel terpilih diperoleh melalui proses ekstraksi dan klasifikasi berbantuan Watase, sehingga hasilnya bergantung pada strategi pencarian, kualitas metadata, dan konfigurasi penyaringan. Selain itu, studi yang ditinjau menggunakan *dataset*, metrik evaluasi, dan protokol *cold-start* yang berbeda sehingga perbandingan statistik langsung antar-model sulit dilakukan. Oleh karena itu, penelitian lanjutan sebaiknya mengembangkan *dataset benchmark* yang secara eksplisit memisahkan *cold-start* pengguna baru dari item baru dan *sparsity* umum, membandingkan model berbasis teks pada split *cold-start* yang sama, serta melaporkan hasil *ablation* untuk memperjelas kontribusi representasi tekstual, fitur sentimen, encoding *transformer*, struktur graf, dan pengetahuan ontologi.

Agenda penelitian lanjutan yang paling mendesak mencakup empat hal yang spesifik. Pertama, penyusunan *dataset benchmark e-commerce* yang memisahkan secara eksplisit pengguna baru tanpa interaksi dari pengguna dengan interaksi minimal dan dari kondisi *sparsity* umum. Kedua, penetapan protokol evaluasi bersama yang menyeragamkan pembagian data, nilai potong K, dan himpunan *baseline* sehingga capaian antar studi dapat dibandingkan secara setara. Ketiga, pelaporan *ablation*

study yang memisahkan kontribusi representasi tekstual, fitur sentimen, encoding *transformer*, struktur graf, dan pengetahuan ontologi. Keempat, pengujian biaya komputasi serta kelayakan penerapan daring, mengingat sebagian model berbasis graf dan model bahasa besar dilaporkan menuntut sumber daya yang besar sehingga berpotensi tidak layak diterapkan pada platform berskala kecil.

Referensi

- [1] D. Nguyen, V. Nguyen, T. Trinh, T. Ho, and H. Le, "A Personalized Product Recommendation Model in e-Commerce based on Retrieval Strategy," *Journal of Open Innovation: Technology, Market, and Complexity*, Vol. 10, No. May, 2024.
- [2] J. David, E. Chapo, and S. Gustavo, "Artificial Intelligence and Recommender Systems in e-Commerce: Trends and Research Agenda," *Intelligent Systems with Applications*, Vol. 24, No. July, 2024, DOI: 10.1016/j.iswa.2024.200435.
- [3] C. Industry, "A Hyper-Personalized Product Recommendation System Focused on Customer Segmentation: An Application in the Fashion Retail Industry," pp. 571-596, 2023.
- [4] H. Yuan and A. A. Hernandez, "User Cold Start Problem in Recommendation Systems: A Systematic Review," *IEEE Access*, Vol. 11, No. November, pp. 136958-136977, 2023, DOI: 10.1109/ACCESS.2023.3338705.
- [5] K. V. Rodpysh, S. J. Mirabedini, and T. Banirostam, "Resolving Cold Start and Sparse Data Challenge in Recommender Systems using Multi-Level Singular Value Decomposition," *Comput. Electr. Eng.*, Vol. 94, No. June, p. 107361, 2021, DOI: 10.1016/j.compeleceng.2021.107361.
- [6] J. Feng, Z. Xia, X. Feng, and J. Peng, "RBPR: A Hybrid Model for the New User Cold Start Problem in Recommender Systems," *Knowledge-Based Syst.*, Vol. 214, p. 106732, 2021, DOI: 10.1016/j.knosys.2020.106732.
- [7] S. S. Roy and A. Kumar, "Metadata and Review-based Hybrid Apparel Recommendation System using Cascaded Large Language Models," No. October, pp. 140053-140071, 2024.
- [8] S. Rama, P. Kollu, and Y. Garapati, "DeepSentRec: A Deep Learning-based Sentiment-Aware Product Recommendation System," pp. 1-26, 2026.
- [9] D. Lim and T. Lee, "Enhanced Recommender System with Sentiment Analysis of Review Text and SBERT Embeddings of Item Descriptions," pp. 1-15, 2026.
- [10] B. M. Aboobaider, "Word Sequential using Deep LSTM and Matrix Factorization to Handle Rating Sparse Data for E-Commerce Recommender System," Vol. 2021, No. 1, 2021, DOI: 10.1155/2021/8751173.
- [11] F. Ikram and H. Farooq, "Semantic Fusion: A Deep Hybrid Multimedia Recommendation System Harnessing Textual and Visual Features," *J. Eng. Res.*, Vol. 14, No. 1, pp. 870-881, 2026, DOI: 10.1016/j.jer.2025.09.004.
- [12] J. J. V. G. Renumol, "Recommender System for Alleviating the Cold-Start Problem," *Educ. Inf. Technol.*, 2021, DOI: 10.1007/s10639-021-10508-0.
- [13] X. Cao, "An Industrial Framework for Cold-Start Recommendation in Few-Shot and Zero-Shot Scenarios," pp. 1-22, 2025.
- [14] M. Park, S. Kim, X. Li, and S. Park, "Enhancing Recommendation with Integration of Extractive and Abstractive Summarization," pp. 1-18, 2026.
- [15] Y. Choi, "Hierarchical Hyperbolic Embeddings for Review-Driven Cross-Domain Recommendation," Vol. 13, No. November, 2025.
- [16] W. Luo, L. Zhang, Y. Cui, and Y. Jin, "Context2Rec: Leveraging Comment Text Semantics and Sequential Features for Enhanced Recommendation Systems," *Pattern Recognit. Lett.*, Vol. 189, No. May 2024, pp. 136-142, 2025, DOI: 10.1016/j.patrec.2025.01.009.
- [17] J. Dong, Z. Shen, H. Luo, T. Lyu, and W. Gao, "Graph Attention and Text Semantics Improve Personalized Recommendation," pp. 1-14, 2026.

- [18] Y. Cui, H. Yu, X. Guo, H. Cao, and L. Wang, "RAKCR: Reviews Sentiment-Aware based Knowledge Graph Convolutional Networks for Personalized Recommendation," *Expert Syst. Appl.*, Vol. 248, No. February, p. 123403, 2024, DOI: 10.1016/j.eswa.2024.123403.
- [19] S. Ray, "Approaches and Algorithms to Mitigate Cold Start Problems in Recommender Systems: A Systematic," *J. Intell. Inf. Syst.*, pp. 341-366, 2022, DOI: 10.1007/s10844-022-00698-5.
- [20] I. Raeesi, V. Laya, M. Seyed, and M. Jafar, *Using Text Mining Algorithms in Identifying Emerging Trends for Recommender Systems*. Springer Netherlands, 2021, DOI: 10.1007/s11135-021-01177-9.
- [21] G. Mustafa et al., "OntoCommerce: Incorporating Ontology and Sequential Pattern Mining for Personalized E-Commerce Recommendations," No. March, pp. 42329-42342, 2024.
- [22] Z. Khan, M. I. Hussain, N. Iltaf, J. Kim, and M. Jeon, "Contextual Recommender System for E-Commerce Applications," *Applied Soft Computing*, Vol. 109, p. 107552, 2021, DOI: 10.1016/j.asoc.2021.107552.
- [23] J. Chambua and Z. Niu, "Review Text based Rating Prediction Approaches: Preference Knowledge Learning, Representation and Utilization," *Artificial Intelligence Review*, Vol. 54, No. 2, pp. 1171-1200, 2021, DOI: 10.1007/s10462-020-09873-y.
- [24] M. Mirzaei, J. Sander, and E. Stroulia, "Personalized and Explainable Aspect-based Recommendation using Latent Opinion Groups," *Computational Intelligence*, 2026.
- [25] K. Saini et al., "A Switching-based Deep Learning Framework for Personalized and Adaptive E-Commerce Recommendations," *Scientific Reports*, 2026.
- [26] R. Esmeli, H. Abdullahi, M. Bader-El-Den, and A. S. Can, "Session Context Data Integration to Address the Cold Start Problem in e-Commerce Recommender Systems," *Decision Support Systems*, 2024.
- [27] T. De Pessemier, B. Willems, and L. Martens, "Active Learning Algorithm for Alleviating the User Cold Start Problem of Recommender Systems," *Scientific Reports*, 2025.
- [28] J. Misztal-Radecka, B. Indurkha, and A. Smywinski-Pohl, "Meta-User2Vec Model for Addressing the User and Item Cold-Start Problem in Recommender Systems," *User Modeling and User-Adapted Interaction*, Vol. 31, pp. 261-286, 2021, DOI: 10.1007/s11257-020-09282-4.
- [29] F. S. Alamri, A. Rehman, B. Alghofaily, A. Ahmed, and K. Saleem, "Leveraging RAG with Transformer for Context-based Personalized Recommendations," *IEEE Access*, 2025, DOI: 10.1109/ACCESS.2025.3574073.
- [30] R. Patel and P. Thakkar, "IBCFaiCDR: Auxiliary Data-Driven Item-based Collaborative Filtering in Cross-Domain RSs to Address User Cold Start Problem," *Results in Engineering*, Vol. 24, p. 103257, 2024.
- [31] N. Hussain et al., "PRUS: Product Recommender System based on User Specifications and Customers Reviews," *IEEE Access*, 2023, DOI: 10.1109/ACCESS.2023.3299818.
- [32] G. Shen, J. Tan, Z. Liu, and X. Kong, "Enhancing Interactive Graph Representation Learning for Review-based Item Recommendation," *Computer Science and Information Systems*, Vol. 19, No. 2, pp. 573-593, 2022, DOI: 10.2298/CSIS210228064S.
- [33] M. Kawai, H. Sato, and T. Shiohama, "Topic Model-based Recommender Systems and Their Applications to Cold-Start Problems," *Expert Systems with Applications*, Vol. 202, p. 117129, 2022.
- [34] V. K. Sejwal and M. Abulaish, "A Hybrid Recommendation Technique using Topic Embedding for Rating Prediction and to Handle Cold-Start Problem," *Expert Systems with Applications*, Vol. 209, p. 118307, 2022.
- [35] S. Gheewala, S. Xu, and S. Yeom, "Deep Shared Learning and Attentive Domain Mapping for Cross-Domain Recommendation," *User Modeling and User-Adapted Interaction*, Vol. 34, pp. 1981-2038, 2024, DOI: 10.1007/s11257-024-09416-y.
- [36] D. Rodriguez-Baena, F. Gomez-Vela, A. Lopez-Fernandez, M. Garcia-Torres, and F. Divina, "BinRec: A Novel UBCF Approach based on Biclustering," *Applied Intelligence*, Vol. 55, 2025, DOI: 10.1007/s10489-025-06725-6.

- [37] S. Kim, L. Jin, Q. Li, X. Li, C. Yoon, and J. Kim, "Utilizing Advanced Transformers for Summarized Review-Aware Recommender System," IEEE Access, 2025, DOI: 10.1109/ACCESS.2025.3598937.
- [38] Z. Li, Z. Liu, G. Sa, M. Hou, J. Sun, and J. Tan, "Talk with Graph: A Fuzzy Sentiment-Aware Framework for Explainable Recommendation," Expert Systems with Applications, 2026, DOI: 10.1016/j.eswa.2025.129707.
- [39] N. A. Osman, S. A. Mohd Noah, M. Darwich, and M. Mohd, "Integrating Contextual Sentiment Analysis in Collaborative Recommender Systems," PLOS ONE, Vol. 16, No. 3, p. e0248695, 2021, DOI: 10.1371/journal.pone.0248695.
- [40] H. Xia and Y. Wang, "Enhancing Neural Collaborative Filtering for Product Recommendation by Integrating Sales Data and User Satisfaction," Electronics, Vol. 14, No. 16, p. 3165, 2025, DOI: 10.3390/electronics14163165.
- [41] G. Diao, C. Li, Q. Liu, and Z. Liu, "Empirical Study on the Application of Deep Learning in User Behavior Prediction and Personalized Recommendation in E-Commerce," Journal of Organizational and End User Computing, Vol. 37, No. 1, 2025, DOI: 10.4018/JOEUC.383512.
- [42] W. Zhang, J. Zhang, and Y. Zhang, "A Comment Aspect-Level User Preference Transfer Model for Cross-Domain Recommendations," Information Resources Management Journal, Vol. 37, No. 1, 2024, DOI: 10.4018/IRMJ.345360.
- [43] Y. Lu, K. Nakamura, and R. Ichise, "HyperRS: Hypernetwork-based Recommender System for the User Cold-Start Problem," IEEE Access, Vol. 11, pp. 5453-5463, 2023, DOI: 10.1109/ACCESS.2023.3236391.
- [44] S. I. H. Jafri, R. Ghazali, I. Javid, Z. Mahmood, and A. A. A. Hassan, "Deep Transfer Learning with Multimodal Embedding to Tackle Cold-Start and Sparsity Issues in Recommendation System," PLOS ONE, Vol. 17, No. 8, p. e0273486, 2022, DOI: 10.1371/journal.pone.0273486.
- [45] L. Gao and J. Li, "E-Commerce Personalized Recommendation Model based on Semantic Sentiment," 2022.
- [46] C. C. Chen, P. L. Lai, and C. Y. Chen, "ColdGAN: An Effective Cold-Start Recommendation System for New Users based on Generative Adversarial Networks," Applied Intelligence, Vol. 53, pp. 8302-8317, 2023, DOI: 10.1007/s10489-022-04005-1.
- [47] Y. Cai, W. Ke, E. Cui, and F. Yu, "A Deep Recommendation Model of Cross-Grained Sentiments of User Reviews and Ratings," Information Processing and Management, Vol. 59, p. 102842, 2022, DOI: 10.1016/j.ipm.2021.102842.
- [48] N. Heidari, M. K. Rafsanjani, and A. Koochari, "An Attention-based Deep Learning Method for Solving the Cold-Start and Sparsity Issues of Recommender Systems," Knowledge-Based Systems, 2022, DOI: 10.1016/j.knosys.2022.109835.
- [49] Q. Y. Shambour, N. M. Turab, and O. Y. Adwan, "An Effective e-Commerce Recommender System based on Trust and Semantic Information," Cybernetics and Information Technologies, Vol. 21, No. 1, pp. 103-118, 2021, DOI: 10.2478/cait-2021-0008.
- [50] C. Yang, X. Chen, L. Liu, and P. Sweetser, "Leveraging Semantic Features for Recommendation: Sentence-Level Emotion Analysis," Information Processing and Management, Vol. 58, p. 102543, 2021, DOI: 10.1016/j.ipm.2021.102543.
- [51] P. Perez-Nunez, P. Buitelaar, J. Diez, O. Luaces, and A. Bahamonde, "Attention-Inspired Text-based Recommender System with Explanatory Capabilities," Applied Soft Computing, Vol. 184, p. 113650, 2025, DOI: 10.1016/j.asoc.2025.113650.
- [52] A. L. Karn et al., "Customer Centric Hybrid Recommendation System for E-Commerce Applications by Integrating Hybrid Sentiment Analysis," Electronic Commerce Research, 2022, DOI: 10.1007/s10660-022-09630-z.
- [53] L.-H. Wu, "BayesSentiRS: Bayesian Sentiment Analysis for Addressing Cold Start and Sparsity in Ranking-based Recommender Systems," Expert Systems with Applications, Vol. 238, p. 121930, 2024, DOI: 10.1016/j.eswa.2023.121930.

- [54] M. M. Bendouch, F. Frasincar, and T. Robal, "A Visual-Semantic Approach for Building Content-based Recommender Systems," *Information Systems*, Vol. 117, p. 102243, 2023, DOI: 10.1016/j.is.2023.102243.
- [55] S. G. K. Patro et al., "A Conscious Cross-Breed Recommendation Approach Confining Cold-Start in Electronic Commerce Systems," *IEEE Access*, 2023, DOI: 10.1109/ACCESS.2023.3274844.
- [56] M. J. Page et al., "The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews," *BMJ*, Vol. 372, p. n71, 2021, DOI: 10.1136/bmj.n71.
- [57] B. Kitchenham, L. Madeyski, and D. Budgen, "SEGRESS: Software Engineering Guidelines for REporting Secondary Studies," *IEEE Transactions on Software Engineering*, Vol. 49, No. 3, pp. 1273-1298, 2023, DOI: 10.1109/TSE.2022.3174092.