

Studi Literatur Information Retrieval System Semantik Untuk Pencarian Produk E-Commerce

Maulana Farras¹, Silvia Hanum², Roberto Kaban³

^{1,2}Fakultas Sains dan Teknologi, Program Studi Rekayasa Perangkat Lunak, Institut Teknologi dan Bisnis Indonesia
Jl. T. Amir Hamzah, Desa Tandem Hilir I, Kec. Hamparan Perak, Kab. Deli Serdang, Sumatera Utara

³Fakultas Sains dan Teknologi, Program Studi Teknik Informatika, Institut Teknologi dan Bisnis Indonesia
Jl. T. Amir Hamzah, Desa Tandem Hilir I, Kec. Hamparan Perak, Kab. Deli Serdang, Sumatera Utara

¹maulanafarras030520@gmail.com, ²hanumsilvia05@gmail.com, ³roberto.kaban@yahoo.com

DOI: <https://doi.org/10.58918/lofian.v5i2.291>

Corresponding Author:

Roberto Kaban
roberto.kaban@yahoo.com

History:

Submitted: 01-02-2026
Accepted: 10-02-2026
Published: 23-02-2026

License:



Abstract

This research aims to analyze the development of the Semantic Information Retrieval System (Semantic IRS) approach in e-commerce product search based on descriptions through a literature study of 15 scientific articles consisting of national and international publications. The results of the analysis show that 33% of articles use the semantic IR and dense retrieval approaches as the basis for semantic mapping between queries and product documents. The late interaction and multimodal semantic retrieval approaches were each applied in 27% of articles, indicating an increasing research focus on token-level semantic interaction modeling and the integration of textual and visual information. Additionally, 13% of articles utilized query expansion and semantic relation modeling as supporting methods to improve search relevance. In terms of methodology, 80% of article used a quantitative experimental approach with information retrieval system metric-based evaluation, and 67% of articles adopted neural models. Overall, these quantitative findings indicate that neural model-based Semantic IR, late interaction, and multimodal approaches are the dominant and most relevant directions for handling long and unstructured description-based product searches in modern e-commerce systems.

Keywords: Semantic Information Retrieval System, Description-Based Product Search, Dense Retrieval, Late Interaction, Multimodal Product Search.

Abstrak

Penelitian ini bertujuan untuk menganalisis perkembangan pendekatan Semantic Information Retrieval System (Semantic IRS) dalam pencarian produk e-commerce berbasis deskripsi melalui studi literatur terhadap 15 artikel ilmiah yang terdiri dari publikasi nasional dan internasional. Hasil analisis menunjukkan bahwa 33% artikel menggunakan pendekatan semantic IR dan dense retrieval sebagai fondasi pemetaan semantik antara kueri dan dokumen produk. Pendekatan late interaction dan multimodal semantic retrieval masing-masing diterapkan pada 27% artikel, menandakan meningkatnya fokus penelitian pada pemodelan interaksi semantik tingkat token serta integrasi informasi teks dan visual. Selain itu, sebanyak 13% artikel memanfaatkan query expansion dan pemodelan relasi semantik sebagai metode pendukung dalam meningkatkan relevansi pencarian. Dari sisi metodologi, 80% artikel menggunakan pendekatan eksperimental kuantitatif dengan evaluasi berbasis metrik information retrieval system, dan 67% artikel mengadopsi model neural. Secara keseluruhan, temuan kuantitatif ini menunjukkan bahwa pendekatan Semantic IR berbasis model neural, late interaction, dan multimodal merupakan arah dominan dan paling relevan untuk menangani pencarian produk berbasis deskripsi yang panjang dan tidak terstruktur pada sistem e-commerce modern.

Kata Kunci: Semantic Information Retrieval System, Pencarian Produk Berbasis Deskripsi, Dense Retrieval, Late Interaction, Multimodal Product Search.

1. Pendahuluan

Perkembangan *e-commerce* dalam beberapa tahun terakhir telah mengalami pertumbuhan yang sangat pesat, baik dari sisi jumlah pengguna, variasi produk, maupun kompleksitas interaksi pengguna dengan sistem digital. Platform *e-commerce* modern tidak lagi hanya berfungsi sebagai media transaksi jual beli, tetapi juga sebagai ekosistem informasi yang menampung jutaan data produk dengan karakteristik, atribut, dan deskripsi yang sangat beragam. Dalam konteks ini, sistem pencarian produk menjadi komponen krusial karena berperan langsung dalam menghubungkan kebutuhan pengguna dengan produk yang tersedia. Ketidakefektifan sistem pencarian dapat menyebabkan kesulitan pengguna dalam menemukan produk yang relevan, yang pada akhirnya berdampak pada menurunnya kepuasan pengguna dan potensi transaksi [1].

Seiring meningkatnya skala data dan heterogenitas deskripsi produk, pola pencarian pengguna di *e-commerce* juga mengalami perubahan signifikan. Pengguna tidak selalu mengetahui nama produk, merek, atau istilah teknis yang tepat, sehingga pencarian sering dilakukan menggunakan deskripsi kebutuhan, fungsi, atau ciri produk secara naratif. Kondisi ini menimbulkan tantangan serius bagi sistem pencarian berbasis kata kunci (*keyword-based search*) yang masih mengandalkan pencocokan leksikal secara literal. Perbedaan istilah, sinonimi, ambiguitas bahasa, serta variasi penulisan menyebabkan hasil pencarian sering kali tidak relevan atau gagal menampilkan produk yang sebenarnya sesuai dengan kebutuhan pengguna [1].

Information Retrieval System (IRS) merupakan bidang keilmuan dalam ilmu komputer yang berfokus pada proses pencarian dan pengambilan informasi relevan dari kumpulan data berskala besar berdasarkan kebutuhan pengguna. Sistem IR menjadi pondasi utama dalam pengembangan mesin pencari, sistem pencarian produk, serta sistem rekomendasi pada platform digital, termasuk *e-commerce*. Model IR klasik seperti Boolean Model, Vector Space Model (VSM), dan BM25 telah digunakan secara luas karena efisiensi dan kemudahan implementasinya. Namun, berbagai studi menunjukkan bahwa model-model tersebut memiliki keterbatasan dalam memahami hubungan makna antar kata, sehingga kurang optimal dalam menangani kueri deskriptif yang bersifat panjang dan tidak terstruktur [2].

Keterbatasan pendekatan leksikal murni mendorong berkembangnya paradigma *Information Retrieval System* semantik, yang bertujuan untuk memahami makna dan konteks di balik teks, bukan sekadar kemunculan kata. IR semantik memanfaatkan teknik pemrosesan bahasa alami (*Natural Language Processing*) serta representasi vektor untuk menangkap hubungan semantik antar istilah. Pendekatan ini memungkinkan sistem untuk mengenali kesamaan makna meskipun istilah yang digunakan berbeda secara tekstual, sehingga lebih adaptif terhadap pola pencarian berbasis deskripsi yang umum terjadi pada platform *e-commerce* modern [2].

Dalam implementasinya pada *e-commerce*, IR semantik memiliki peran strategis dalam meningkatkan kualitas pengalaman pencarian produk. Sistem yang mampu memahami konteks dan maksud pencarian pengguna dapat menyajikan hasil yang lebih relevan, mengurangi friksi pencarian, serta membantu pengguna dalam proses pengambilan keputusan. Hal ini menjadi semakin penting mengingat fenomena *information overload* yang muncul akibat melimpahnya pilihan produk pada platform *e-commerce*. Tanpa mekanisme pencarian yang cerdas, pengguna berpotensi mengalami kebingungan dan kelelahan kognitif dalam memilih produk yang sesuai [3].

Selain sistem pencarian, pendekatan semantik juga banyak diterapkan pada sistem rekomendasi produk berbasis konten (*content-based filtering*). Pendekatan ini memanfaatkan kesamaan atribut produk, seperti nama dan deskripsi, untuk merekomendasikan produk yang relevan kepada pengguna. Penelitian menunjukkan bahwa representasi teks menggunakan teknik seperti TF-IDF dan *cosine similarity* mampu memberikan rekomendasi yang cukup efektif, khususnya ketika data interaksi pengguna terbatas. Namun, pendekatan ini masih menghadapi keterbatasan dalam menangkap makna kontekstual yang lebih dalam, sehingga memerlukan pengayaan semantik agar hasil yang dihasilkan semakin relevan [4].

Perkembangan terkini dalam IR semantik juga menunjukkan pergeseran menuju pendekatan berbasis model *neural dan embedding* semantik yang lebih canggih. Model-model ini mampu memetakan teks ke dalam ruang vektor berdimensi tinggi yang merepresentasikan makna laten, sehingga hubungan semantik antar kueri dan dokumen dapat diukur secara lebih akurat. Dalam konteks *e-commerce*, pendekatan ini terbukti efektif dalam menangani kueri ambigu, *long-tail query*, serta variasi bahasa yang sering digunakan pengguna dalam mendeskripsikan kebutuhan produk [2].

Selain aspek tekstual, sejumlah penelitian juga menekankan pentingnya integrasi informasi multimodal dalam sistem pencarian dan rekomendasi *e-commerce*. Produk tidak hanya direpresentasikan melalui deskripsi teks, tetapi juga melalui gambar yang mengandung informasi visual penting seperti warna, bentuk, dan gaya. Pendekatan

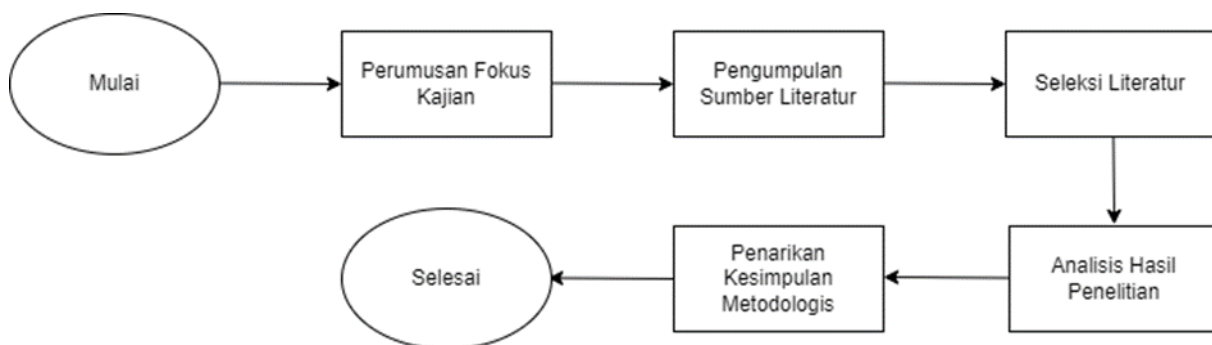
multimodal yang mengombinasikan teks dan gambar terbukti mampu meningkatkan pemahaman sistem terhadap karakteristik produk, terutama pada skenario pencarian yang kompleks dan ambigu. Implementasi *Retrieval-Augmented Generation* (RAG) multimodal menunjukkan potensi besar dalam meningkatkan kualitas pencarian dan rekomendasi produk secara holistik [5].

Di sisi lain, pengembangan fitur pencarian dan filter produk pada aplikasi *e-commerce* juga menunjukkan bahwa kebutuhan pengguna sangat beragam dan kontekstual. Sistem pencarian yang baik tidak hanya harus akurat, tetapi juga mudah digunakan dan responsif terhadap kebutuhan pengguna. Penelitian terkait pengembangan fitur pencarian produk pada aplikasi *e-commerce* menegaskan bahwa optimalisasi mekanisme pencarian berkontribusi langsung terhadap efisiensi operasional dan kepuasan pengguna, terutama pada platform dengan skala produk yang besar [6].

Berdasarkan berbagai kajian tersebut, dapat disimpulkan bahwa sistem *Information Retrieval System* semantik memiliki peran yang sangat penting dalam menjawab tantangan pencarian produk *e-commerce* berbasis deskripsi. Meskipun telah banyak pendekatan dan teknik yang dikembangkan, masih diperlukan pemahaman komprehensif mengenai model, teknik, serta metode evaluasi yang digunakan dalam IR semantik agar implementasinya dapat disesuaikan dengan kebutuhan domain *e-commerce* secara optimal. Oleh karena itu, penelitian ini disusun dalam bentuk studi literatur untuk mengkaji secara sistematis berbagai model *Information Retrieval System* semantik, teknik pendukung, serta metode evaluasi yang relevan dalam konteks pencarian produk *e-commerce* berdasarkan deskripsi, sehingga dapat menjadi landasan teoretis dan praktis bagi pengembangan sistem pencarian yang lebih efektif dan adaptif di masa mendatang.

2. Metodologi Penelitian

Penelitian ini menggunakan pendekatan **studi literatur (literature study)** dengan tujuan untuk mengkaji, menganalisis hasil-hasil penelitian terdahulu yang berkaitan dengan sistem *Information Retrieval System* (IRS) semantik dalam konteks pencarian produk *e-commerce* berbasis deskripsi. Studi literatur dipilih karena mampu memberikan gambaran komprehensif mengenai perkembangan metode, teknik, serta hasil evaluasi yang telah dilakukan oleh peneliti sebelumnya, sehingga memungkinkan penarikan kesimpulan yang objektif mengenai metode IR yang paling efektif dan relevan untuk permasalahan yang dikaji.



Gbr. 1. Alur Metodologi Penelitian

2.1. Tahapan Metode Penelitian

Untuk mencapai tujuan penelitian, metode studi literatur ini dilaksanakan melalui beberapa tahapan sistematis sebagai berikut:

a. Perumusan Fokus Kajian

Fokus penelitian ditetapkan pada kajian sistem *Information Retrieval System* semantik yang diterapkan untuk pencarian produk *e-commerce* berbasis deskripsi. Fokus ini diarahkan pada kondisi ketika pengguna tidak menyebutkan nama produk secara eksplisit, melainkan menggunakan deskripsi kebutuhan, fungsi, atau karakteristik produk. Berdasarkan fokus tersebut, ruang lingkup kajian mencakup analisis model IR yang digunakan, teknik pemrosesan dan representasi teks, serta pendekatan evaluasi kinerja yang diterapkan dalam penelitian sebelumnya.

b. Pengumpulan Sumber Literatur

Sumber data penelitian diperoleh dari 27 artikel ilmiah yang relevan dengan topik *Information Retrieval System* semantik dan pencarian produk berbasis deskripsi. Artikel-artikel tersebut berasal dari jurnal nasional dan internasional yang terdiri dari ACM Digital Library, Elsevier, Scopus, ECNLP WWW Companion, Sinta 2 hingga Sinta 5. Pemilihan sumber dilakukan secara selektif dengan mempertimbangkan kesesuaian topik, kontribusi ilmiah, dan keterkaitan langsung dengan permasalahan pencarian produk pada platform *e-commerce*.

c. Seleksi Literatur

Setiap artikel yang telah dikumpulkan kemudian diseleksi dan dianalisis sehingga menjadi 15 artikel untuk mengidentifikasi karakteristik penelitian, meliputi tujuan penelitian, jenis data yang digunakan, model *Information Retrieval System* yang diterapkan, serta teknik pemrosesan teks yang digunakan

d. Analisis Hasil Penelitian

Literatur yang telah diseleksi dianalisis secara komparatif untuk menilai keunggulan dan keterbatasan masing-masing metode IR dalam mendukung pencarian produk *e-commerce* berbasis deskripsi. Analisis difokuskan pada identifikasi jenis metode semantic IR, meliputi representasi vektor dan semantik laten, *collaborative* dan *content-based filtering*, analisis teks dan relasi semantik, pemodelan berbasis graf, serta model neural dan multimodal. Hasil dari setiap kelompok metode dianalisis secara deskriptif dan komparatif untuk melihat kecenderungan penggunaan metode, kelebihan dan keterbatasannya, serta relevansinya dalam mendukung pencarian produk *e-commerce* berbasis deskripsi.

e. Penarikan Kesimpulan Metodologis

Berdasarkan hasil analisis literatur, dilakukan penarikan kesimpulan untuk menentukan metode *Information Retrieval System* semantik yang paling sesuai dan efektif dalam pencarian produk *e-commerce* berbasis deskripsi. Penentuan metode terbaik tidak hanya didasarkan pada nilai akurasi tertinggi, tetapi juga mempertimbangkan konsistensi performa, fleksibilitas metode dalam menangani variasi bahasa, serta kesesuaian dengan kebutuhan implementasi sistem *e-commerce*.

2.2. Seleksi Artikel Penelitian

Pada tahap awal penelitian, proses pengumpulan literatur menghasilkan sebanyak 27 artikel ilmiah yang relevan dengan topik *Information Retrieval System* semantik dan pencarian produk pada platform *e-commerce*. Namun, tidak seluruh artikel tersebut dapat digunakan sebagai objek utama studi literatur karena adanya perbedaan tingkat relevansi dan kesesuaian dengan fokus penelitian. Oleh karena itu, dilakukan proses seleksi bertahap untuk memastikan bahwa artikel yang dianalisis benar-benar sesuai dengan kriteria pencarian produk berbasis deskripsi serta memiliki kebaruan yang memadai. Proses seleksi ini dilakukan melalui dua tahap utama, yaitu seleksi berdasarkan relevansi topik dan seleksi berdasarkan tahun publikasi.

Tabel 1

Tahapan Seleksi Artikel dalam Studi Literatur

| Kriteria / Tahap Seleksi | Artikel Awal | Seleksi Tahap I | Seleksi Tahap I | Seleksi Tahap II | Artikel Akhir |
|--------------------------|-----------------------------|--|-------------------------|------------------------|---------------------------------|
| Jumlah Artikel | 27 | 2 dieleminasi | 25 | 10 dieleminasi | 15 |
| Dasar Seleksi | Hasil pengumpulan literatur | Tidak relevan dengan pencarian produk berbasis deskripsi | Lolos seleksi relevansi | Tahun publikasi < 2021 | Digunakan dalam studi literatur |

2.3. Distribusi Fokus Penelitian pada Literatur yang Dianalisis

Berdasarkan pengelompokan terhadap lima belas artikel yang digunakan dalam studi literatur ini, fokus penelitian menunjukkan kecenderungan yang cukup jelas pada pengembangan *Information Retrieval System* semantik untuk pencarian produk *e-commerce*. Secara umum, topik penelitian dapat diklasifikasikan ke dalam empat kelompok besar, yaitu *semantic/dense retrieval*, *late interaction*, *multimodal retrieval*, serta *query understanding* dan *personalization*.

Tabel 2

Distribusi Fokus Penelitian pada 15 Artikel

| No | Fokus Penelitian Utama | Jumlah Artikel | Persentase |
|-------|---|----------------|------------|
| 1 | Semantic IR & Dense Retrieval | 5 | 33% |
| 2 | Late Interaction untuk Product Search | 4 | 27% |
| 3 | Multimodal Semantic Retrieval (Teks + Gambar) | 4 | 27% |
| 4 | Query Expansion, Similarity & Personalisasi | 2 | 13% |
| Total | | 15 | 100% |

Berdasarkan Tabel 1, dapat diamati bahwa fokus penelitian dalam lima belas artikel yang dianalisis didominasi oleh pendekatan *semantic information retrieval* dan *dense retrieval*, dengan proporsi sebesar 33%. Dominasi ini menunjukkan bahwa pemetaan kueri dan dokumen produk ke dalam ruang vektor semantik masih menjadi fondasi utama dalam pengembangan sistem pencarian produk e-commerce, khususnya untuk mengatasi keterbatasan pencarian berbasis kata kunci pada kueri deskriptif dan *long-tail*. Pendekatan ini banyak dipilih karena efisiensinya dalam skala besar serta kemampuannya menangkap kesamaan makna secara laten.

Selain itu, pendekatan *late interaction* dan *multimodal semantic retrieval* masing-masing menempati proporsi yang signifikan, yaitu sebesar 27%. Hal ini mengindikasikan adanya pergeseran fokus penelitian menuju model yang mampu mempertahankan detail semantik tingkat token serta mengintegrasikan informasi visual dan tekstual produk. Sementara itu, penelitian yang berfokus pada *query expansion*, *query similarity*, dan personalisasi masih berada pada proporsi yang lebih kecil. Temuan ini menunjukkan bahwa meskipun aspek pemahaman kueri terus berkembang, penelitian terkini lebih menekankan pada peningkatan representasi dan mekanisme pencocokan semantik sebagai solusi utama dalam pencarian produk e-commerce berbasis deskripsi.

2.4. Status Indeksasi dan Reputasi Publikasi

Kualitas literatur dalam studi ini ditinjau melalui status indeksasi jurnal atau prosiding tempat artikel diterbitkan. Penilaian ini penting untuk memastikan bahwa studi literatur didasarkan pada sumber yang kredibel dan diakui secara akademik.

Tabel 3

Klasifikasi Indeksasi dan Rumah Publikasi Literatur

| No | Kategori Publikasi | Indeksasi / Rumah Publikasi | Jumlah Artikel |
|-------|--|--|----------------|
| 1 | Jurnal / Prosiding Internasional Bereputasi Tinggi | ACM Digital Library (SIGIR, WWW, KDD, ECNLP, Workshop) | 6 |
| 2 | Jurnal Internasional Bereputasi | Elsevier (Scopus Q1) | 1 |
| 3 | Workshop Internasional Bereputasi | ECNLP, WWW Companion | 3 |
| 4 | Jurnal Nasional Terakreditasi SINTA 2–3 | SINTA 2, SINTA 3 | 2 |
| 5 | Jurnal Nasional Terakreditasi SINTA 4–5 | SINTA 4, SINTA 5 | 3 |
| Total | | | 15 |

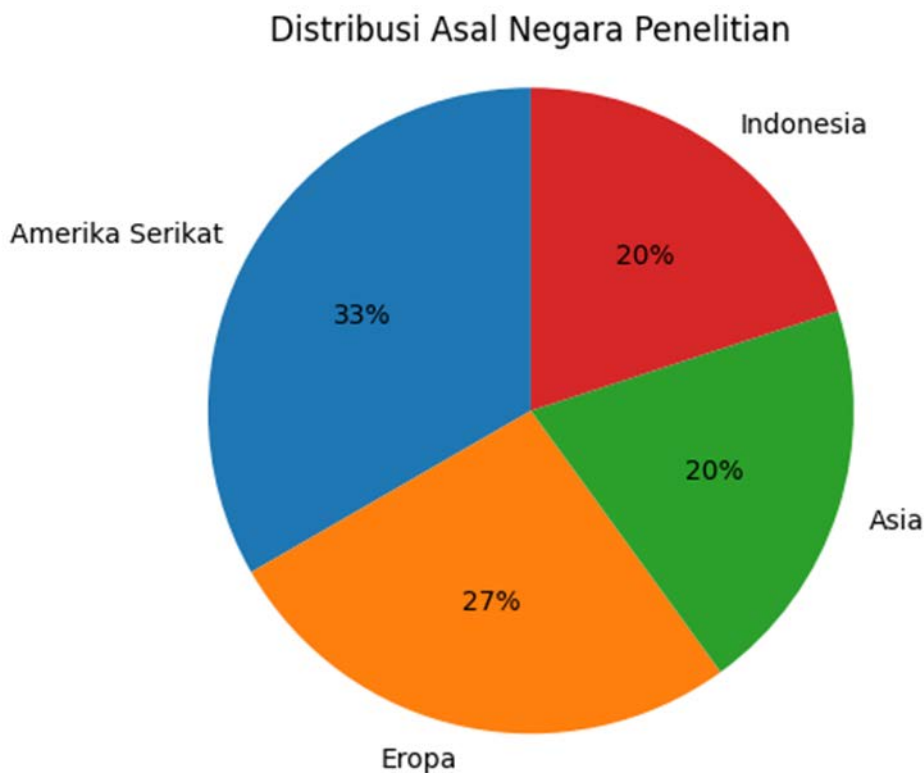
Dari Tabel 2, dapat dilihat bahwa mayoritas sumber literatur dalam studi ini berasal dari publikasi internasional bereputasi, khususnya yang berada di bawah naungan ACM Digital Library dan jurnal internasional terindeks Scopus. Dominasi publikasi ACM menunjukkan bahwa topik *Information Retrieval System* semantik dan pencarian produk e-commerce merupakan fokus utama komunitas riset global di bidang *Information Retrieval*, *Data Mining*, dan *Natural Language Processing*. Konferensi dan workshop seperti SIGIR, WWW, dan KDD dikenal memiliki standar seleksi yang ketat, sehingga artikel-artikel yang diterbitkan mencerminkan kontribusi metodologis dan eksperimental yang signifikan.

Di sisi lain, keberadaan jurnal nasional terakreditasi SINTA 2 hingga SINTA 5 menunjukkan bahwa kajian IR semantik juga berkembang secara aktif dalam konteks penelitian nasional. Artikel-artikel nasional umumnya menekankan aspek konseptual, studi literatur, serta implementasi dan adaptasi metode IR semantik pada konteks lokal. Kombinasi antara literatur internasional bereputasi dan jurnal nasional terakreditasi ini memperkuat landasan

teoritis dan empiris penelitian, sekaligus memastikan bahwa analisis yang dilakukan bersifat komprehensif, kontekstual, dan relevan baik pada skala global maupun nasional.

2.5. Asal Negara Konteks Penelitian

Analisis geografis dilakukan untuk melihat sebaran kontribusi penelitian berdasarkan asal institusi penulis. Hasil analisis menunjukkan bahwa penelitian IR semantik untuk e-commerce bersifat lintas negara dan didominasi oleh negara dengan ekosistem e-commerce dan riset AI yang matang.



Gbr. 2. Distribusi Asal Negara Penelitian

Berdasarkan grafik lingkaran yang disajikan, dapat dilihat bahwa kontribusi penelitian dalam kajian *Information Retrieval System* semantik untuk pencarian produk e-commerce didominasi oleh artikel yang berasal dari Amerika Serikat, dengan proporsi sekitar 33%. Dominasi ini menunjukkan kuatnya peran institusi akademik dan industri teknologi di Amerika Serikat dalam pengembangan model IR berbasis neural, khususnya yang diterapkan pada skala industri e-commerce besar.

Kontribusi dari kawasan Eropa (Belanda, Italia, Spanyol, Inggris) menempati posisi kedua dengan proporsi sekitar 27%, diikuti oleh Asia (Tiongkok, Singapura) dan Indonesia yang masing-masing berkontribusi sekitar 20%. Kehadiran penelitian dari Asia dan Indonesia menunjukkan bahwa kajian IR semantik tidak hanya berkembang di negara dengan ekosistem e-commerce mapan, tetapi juga semakin mendapat perhatian di negara berkembang. Secara keseluruhan, sebaran geografis ini menegaskan bahwa penelitian IR semantik bersifat global dan lintas konteks, serta memperkaya sudut pandang metodologis dan implementatif dalam studi literatur ini.

2.6. Karakteristik Metodologi dan Evaluasi yang Digunakan

Ditinjau dari metodologi, hampir seluruh artikel menggunakan pendekatan eksperimental kuantitatif dengan evaluasi kinerja sistem. Namun, terdapat variasi signifikan pada jenis model dan teknik evaluasi yang digunakan. Pendekatan eksperimental kuantitatif dipilih karena penelitian *Information Retrieval System* pada umumnya menuntut pembuktian empiris melalui pengujian sistem terhadap dataset tertentu serta pengukuran kinerja berbasis metrik yang terstandarisasi.

Lebih lanjut, variasi metodologi yang muncul tidak hanya terletak pada perbedaan jenis model IR yang digunakan, tetapi juga pada cara peneliti membandingkan performa model, mengolah data pelatihan, serta merancang skenario evaluasi. Beberapa artikel berfokus pada pengembangan satu arsitektur model tertentu, sementara artikel lainnya secara eksplisit melakukan perbandingan antara model IR tradisional, *dense retrieval*, hingga *late interaction*. Selain itu, berkembang pula penelitian yang mengintegrasikan data multimodal dan memanfaatkan teknik pembelajaran mendalam berbasis *transformer* untuk meningkatkan pemahaman semantik kueri dan dokumen produk. Oleh karena itu, pemetaan karakteristik metodologi ini penting untuk memberikan gambaran menyeluruh mengenai arah dan kecenderungan riset IR semantik pada pencarian produk e-commerce.

Tabel 4

Karakteristik Metodologi Penelitian pada Literatur

| No | Karakteristik Metodologi | Jumlah Artikel | Penjelasan Karakteristik |
|----|--------------------------------------|----------------|---|
| 1 | Pendekatan Eksperimental Kuantitatif | 12 | Artikel menguji model IR melalui eksperimen sistematis menggunakan dataset tertentu dan mengukur kinerja secara numerik |
| 2 | Model Neural Berbasis Transformer | 10 | Menggunakan arsitektur transformer (BERT, RoBERTa, DistilBERT) untuk representasi kueri dan produk |
| 3 | Dense Retrieval / Dual Encoder | 6 | Kueri dan dokumen direpresentasikan sebagai vektor tunggal dan dicocokkan menggunakan vector similarity |
| 4 | Late Interaction | 4 | Mempertahankan interaksi tingkat token antara kueri dan dokumen untuk menangkap detail semantik |
| 5 | Multimodal Learning (Teks + Gambar) | 4 | Menggabungkan informasi visual dan tekstual produk untuk meningkatkan relevansi pencarian |
| 6 | Studi Komparatif Model IR | 8 | Membandingkan performa beberapa pendekatan IR dalam satu penelitian |
| 7 | Evaluasi dengan Metrik IR Standar | 13 | Menggunakan Precision, Recall, NDCG, MRR untuk menilai kualitas hasil pencarian |
| 8 | Studi Literatur / Konseptual | 3 | Tidak melakukan eksperimen sistem, tetapi menganalisis dan mensintesis penelitian sebelumnya |

Berdasarkan Tabel 3, terlihat bahwa mayoritas artikel menggunakan pendekatan eksperimental kuantitatif dengan memanfaatkan model neural berbasis *transformer*. Hal ini menunjukkan bahwa riset IR semantik untuk pencarian produk e-commerce telah berada pada tahap pengujian empiris yang matang, dengan fokus utama pada peningkatan performa sistem melalui inovasi arsitektur model dan strategi pelatihan.

Selain itu, meskipun pendekatan *dense retrieval* masih banyak digunakan karena efisiensinya, pendekatan *late interaction* mulai mendapatkan perhatian khusus karena kemampuannya mempertahankan detail semantik tingkat token. Meningkatnya jumlah penelitian *multimodal retrieval* juga menandakan adanya pergeseran paradigma dari pencarian berbasis teks menuju pemahaman produk yang lebih komprehensif. Secara keseluruhan, karakteristik metodologi yang beragam ini menunjukkan bahwa penelitian IR semantik berkembang ke arah sistem yang lebih kompleks, adaptif, dan sesuai dengan kebutuhan pencarian produk e-commerce berbasis deskripsi.

3. Hasil Dan Pembahasan

Pada bagian ini menyajikan analisis komprehensif terhadap 15 artikel ilmiah yang dijadikan objek dalam studi literatur ini. Seluruh artikel dianalisis secara sistematis untuk mengkaji bagaimana pendekatan *Semantic Information Retrieval System* (Semantic IRS) dikembangkan dan diterapkan dalam konteks pencarian informasi, khususnya pada pencarian produk berbasis deskripsi di lingkungan e-commerce dan sistem temu kembali informasi kontemporer. Analisis difokuskan pada identifikasi metode semantic IR yang digunakan, konteks implementasi masing-masing penelitian, serta kontribusi setiap artikel dalam meningkatkan relevansi dan kualitas hasil pencarian.

Hasil telaah menunjukkan bahwa penerapan semantic IR pada 15 artikel tidak bersifat homogen, melainkan berkembang melalui beragam pendekatan metodologis. Pendekatan tersebut mencakup representasi vektor dan latent semantic, pemodelan relasi semantik berbasis graf, hingga pemanfaatan arsitektur neural dan teknik multimodal untuk menangkap makna deskriptif secara lebih mendalam. Oleh karena itu, tabel ringkasan artikel

yang disajikan pada bagian berikut berfungsi sebagai hasil pemetaan awal literatur, yang menjadi landasan analitis untuk pembahasan lebih rinci mengenai karakteristik, kelebihan, dan keterbatasan masing-masing pendekatan semantic IR pada subbab selanjutnya.

Tabel 5
 15 Artikel Ilmiah Sebagai Bahan Studi Literatur

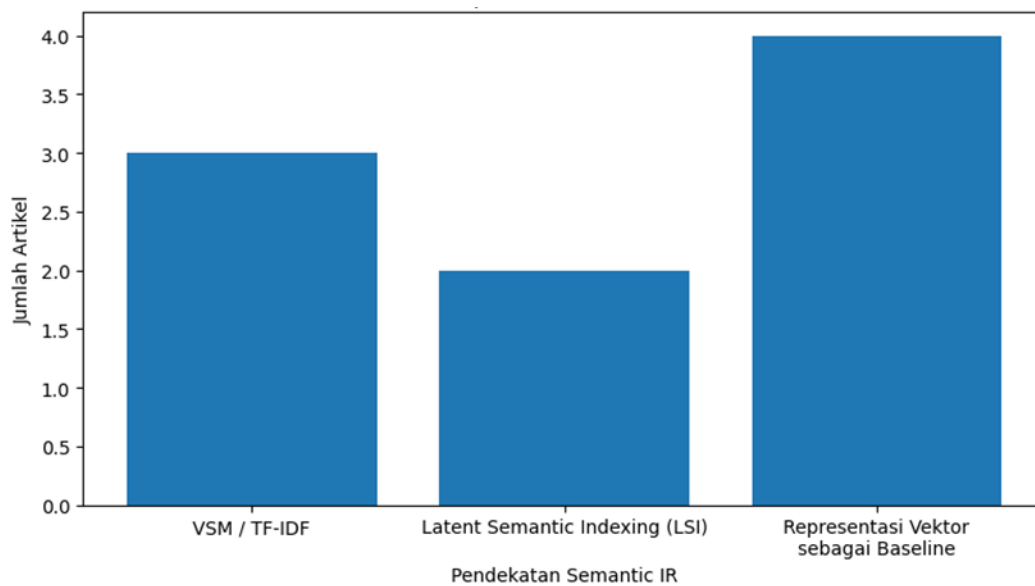
| Judul | Penulis | Tahun Terbit |
|--|---------|--------------|
| Sistem rekomendasi produk <i>e-commerce</i> menggunakan <i>collaborative filtering</i> dan <i>content-based filtering</i> | [7] | 2025 |
| Sistematik literatur sistem temu kembali informasi dengan <i>Vector Space Model</i> dan <i>Depth First Search</i> | [8] | 2021 |
| Pengembangan sistem rekomendasi produk <i>e-commerce</i> menggunakan algoritma <i>collaborative filtering</i> | [9] | 2025 |
| <i>Sentiment analysis of the top 5 e-commerce platforms in Indonesia using text mining and natural language processing (NLP)</i> | [10] | 2023 |
| Implementasi <i>information retrieval system</i> menggunakan metode <i>latent semantic indexing</i> | [11] | 2023 |
| <i>Improving multi-hop knowledge base question answering by learning intermediate supervision signals</i> | [12] | 2021 |
| <i>Modeling user behavior with graph convolution for personalized product search</i> | [13] | 2022 |
| <i>E-commerce search via content collaborative graph neural network</i> | [14] | 2023 |
| <i>Semantic-enhanced modality-asymmetric retrieval for online e-commerce search</i> | [15] | 2023 |
| <i>Evaluation of semantic relations impact in query expansion-based retrieval systems</i> | [16] | 2024 |
| <i>Multimodal Semantic Retrieval for Product Search</i> | [17] | 2025 |
| <i>Multimodal Learning with Online Text Cleaning for E-commerce Product Search</i> | [18] | 2024 |
| <i>Cluster Language Model for Improved E-Commerce Retrieval and Ranking</i> | [19] | 2024 |
| <i>Towards Generalizable Semantic Product Search by Text Similarity Pre-training on Search Click Logs</i> | [20] | 2022 |

3.1. Analisis Metode Semantic Information Retrieval System pada Pencarian Produk Berbasis Deskripsi

3.1.1. Semantic IR Berbasis Representasi Vektor dan Laten

Pendekatan semantic IR berbasis representasi vektor dan semantik laten ditemukan secara dominan pada artikel nasional dan sebagian artikel internasional. Artikel oleh Amalia et al. (2021) melakukan kajian sistematik terhadap penerapan *Vector Space Model* (VSM) dalam sistem temu kembali informasi, yang menunjukkan bahwa representasi vektor mampu menangkap kedekatan semantik dasar antar dokumen meskipun masih bergantung pada kemunculan kata. Pendekatan serupa juga menjadi fondasi dalam penelitian Kalimantan dan Agustin (2023) yang menerapkan *Latent Semantic Indexing* (LSI) untuk memahami keterkaitan makna antara query deskriptif dan dokumen [8], [11].

Keunggulan pendekatan ini terletak pada kesederhanaan dan efisiensi komputasi, sehingga mudah diimplementasikan pada sistem pencarian berskala kecil hingga menengah. Namun, keterbatasan utama yang diidentifikasi dalam kedua artikel tersebut adalah ketidakmampuan model untuk menangkap konteks semantik yang kompleks, terutama pada deskripsi produk yang panjang dan multi-atribut. Keterbatasan ini juga tercermin dalam artikel nasional lain yang masih berfokus pada metode statistik dan *keyword-based matching* sebagai dasar pencarian [8], [11].



Gbr. 3. Distribusi Pendekatan Semantic IR Berbasis Representasi Vektor dan Laten

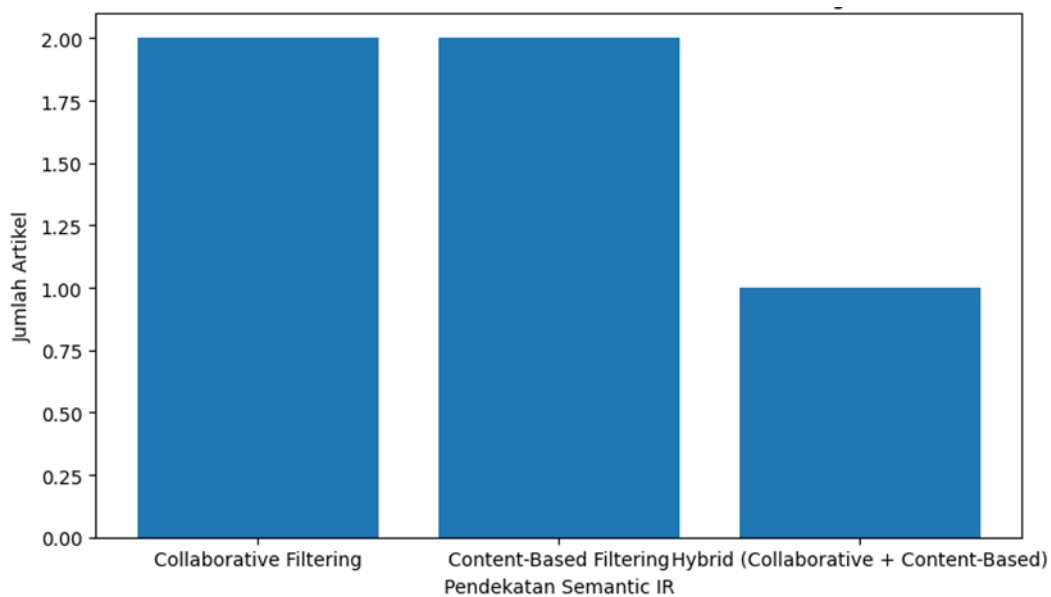
Berdasarkan Gambar 3, terlihat bahwa penggunaan representasi vektor sebagai pendekatan dasar (*baseline*) merupakan yang paling banyak ditemukan dalam artikel yang dianalisis. Hal ini menunjukkan bahwa representasi vektor masih menjadi fondasi utama dalam pengembangan sistem semantic IR, khususnya pada penelitian yang berada pada tahap awal pengembangan atau sebagai pembandingan terhadap metode yang lebih kompleks. Pendekatan *Vector Space Model* dan TF-IDF juga menunjukkan proporsi yang cukup signifikan, yang mengindikasikan bahwa metode tersebut masih relevan untuk pencarian berbasis teks dengan kompleksitas semantik yang relatif rendah.

Sementara itu, pendekatan *Latent Semantic Indexing* muncul dalam jumlah yang lebih terbatas. Meskipun demikian, keberadaannya mencerminkan upaya untuk menangkap hubungan makna laten yang tidak dapat diperoleh melalui pencocokan kata secara langsung. Secara keseluruhan, grafik ini menegaskan bahwa pendekatan semantic IR berbasis representasi vektor dan laten berperan penting sebagai landasan metodologis, namun memiliki keterbatasan dalam menangani pencarian produk berbasis deskripsi yang kompleks, sehingga mendorong berkembangnya pendekatan semantic IR yang lebih kontekstual pada penelitian selanjutnya.

3.1.2. Semantic IR Berbasis Collaborative dan Content-Based Filtering

Pendekatan semantic IR dalam konteks rekomendasi dan pencarian produk e-commerce banyak dibahas dalam artikel Rachmaniar et al. (2025) dan Ramadhani et al. (2025). Kedua artikel ini memanfaatkan *collaborative filtering* dan *content-based filtering* untuk memahami kesamaan semantik antar produk berdasarkan perilaku pengguna dan atribut konten [7], [9].

Meskipun tidak secara eksplisit menggunakan istilah *semantic IR*, pendekatan ini secara konseptual termasuk dalam semantic retrieval karena memanfaatkan relasi laten antara pengguna, produk, dan deskripsi produk. Keunggulan metode ini adalah kemampuannya dalam menangkap preferensi implisit pengguna, sehingga relevan untuk pencarian produk berbasis deskripsi yang bersifat personal. Namun, kelemahannya terletak pada ketergantungan terhadap data historis pengguna serta keterbatasan dalam menangani produk baru (*cold start problem*) [7], [9].



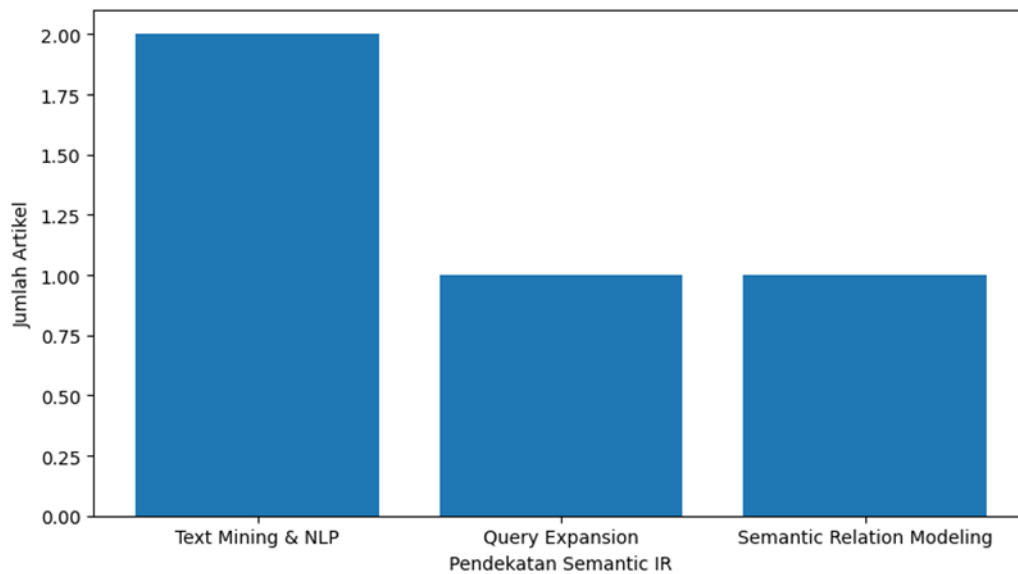
Gbr. 4. Distribusi Artikel dengan Pendekatan Semantic IR Berbasis Collaborative dan Content-Based Filtering

Berdasarkan Gambar 4, terlihat bahwa pendekatan *Collaborative Filtering* dan *Content-Based Filtering* digunakan secara relatif seimbang dalam artikel yang dianalisis. Hal ini menunjukkan bahwa kedua pendekatan tersebut masih menjadi metode yang relevan dalam konteks pencarian dan rekomendasi produk e-commerce, khususnya untuk memahami preferensi pengguna dan kesamaan atribut produk berdasarkan data historis maupun deskripsi konten.

Pendekatan *Collaborative Filtering* umumnya dimanfaatkan untuk menangkap pola kesamaan antar pengguna berdasarkan interaksi sebelumnya, sehingga efektif dalam mendukung pencarian produk yang bersifat personal. Sementara itu, *Content-Based Filtering* lebih menekankan pada kesesuaian antara deskripsi produk dan kebutuhan pengguna, sehingga lebih stabil dalam menangani variasi deskripsi produk. Grafik ini juga menunjukkan bahwa pendekatan hibrida mulai diterapkan, meskipun masih dalam jumlah terbatas, sebagai upaya untuk mengombinasikan kelebihan kedua metode dan meminimalkan keterbatasan seperti *cold-start problem*. Temuan ini sejalan dengan pembahasan pada subbab sebelumnya, yang menempatkan pendekatan collaborative dan content-based sebagai bagian dari evolusi semantic IR menuju sistem pencarian produk berbasis deskripsi yang lebih adaptif.

3.1.3. Semantic IR Berbasis Analisis Teks dan Relasi Semantik

Artikel Sapanji et al. (2023) dan Massai (2024) menyoroti peran analisis teks dan relasi semantik dalam meningkatkan kualitas pencarian dan pemahaman query. Sapanji et al. (2023) menunjukkan bahwa teknik text mining dan natural language processing mampu mengekstraksi makna dan sentimen dari teks dalam konteks e-commerce, yang secara tidak langsung memperkaya pemahaman semantik sistem pencarian [10], [16].



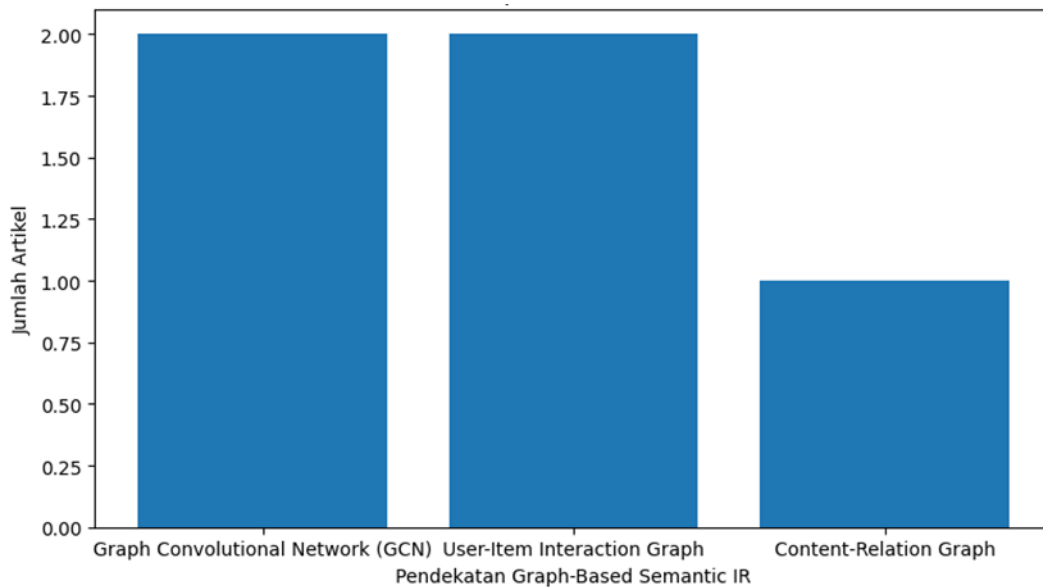
Gbr. 5. Distribusi Artikel dengan Pendekatan Semantic IR Berbasis Analisis Teks dan Relasi Semantik

Berdasarkan Gambar 5, terlihat bahwa pendekatan berbasis *text mining* dan NLP merupakan metode yang paling banyak digunakan dalam kelompok ini. Hal ini menunjukkan bahwa analisis linguistik terhadap teks masih menjadi fondasi penting dalam upaya memahami makna dan konteks kueri serta dokumen, terutama pada pencarian produk berbasis deskripsi yang mengandalkan informasi tekstual. Pendekatan ini umumnya digunakan untuk mengekstraksi fitur semantik, sentimen, serta pola kata yang relevan dari deskripsi produk dan ulasan pengguna.

Sementara itu, pendekatan *query expansion* dan pemodelan relasi semantik muncul dalam jumlah yang lebih terbatas. Meskipun demikian, keberadaan kedua pendekatan tersebut mencerminkan upaya untuk memperkaya representasi kueri dengan menambahkan istilah atau relasi makna yang relevan, sehingga sistem pencarian mampu mengurangi ambiguitas dan meningkatkan relevansi hasil pencarian. Secara keseluruhan, grafik ini menegaskan bahwa analisis teks dan relasi semantik berperan sebagai komponen pendukung yang penting dalam semantic IR, khususnya sebagai tahap penguatan pemahaman kueri sebelum diterapkan pada metode pencarian yang lebih kompleks.

3.1.4. *Semantic IR Berbasis Graph dan Pemodelan Relasi*

Pendekatan berbasis graf dibahas secara mendalam dalam artikel Fan et al. (2022) dan Xv et al. (2023). Kedua artikel tersebut memanfaatkan *graph convolution* untuk memodelkan relasi antara pengguna, produk, dan konten. Pendekatan ini memungkinkan sistem pencarian memahami struktur relasi yang kompleks, yang sulit ditangkap oleh model vektor konvensional [13], [14].



Gbr. 6. Distribusi Artikel dengan Pendekatan Semantic IR Berbasis Graph dan Pemodelan Relasi

Berdasarkan Gambar 6, terlihat bahwa pendekatan *Graph Convolutional Network* dan *user-item interaction graph* merupakan metode yang paling banyak digunakan dalam kelompok *graph-based semantic IR*. Hal ini menunjukkan bahwa pemodelan hubungan struktural antar entitas, seperti pengguna, produk, dan konten, dipandang efektif dalam menangkap keterkaitan semantik yang tidak dapat direpresentasikan secara optimal melalui model vektor konvensional. Pendekatan ini sangat relevan dalam konteks pencarian produk e-commerce yang melibatkan interaksi kompleks antara pengguna dan item.

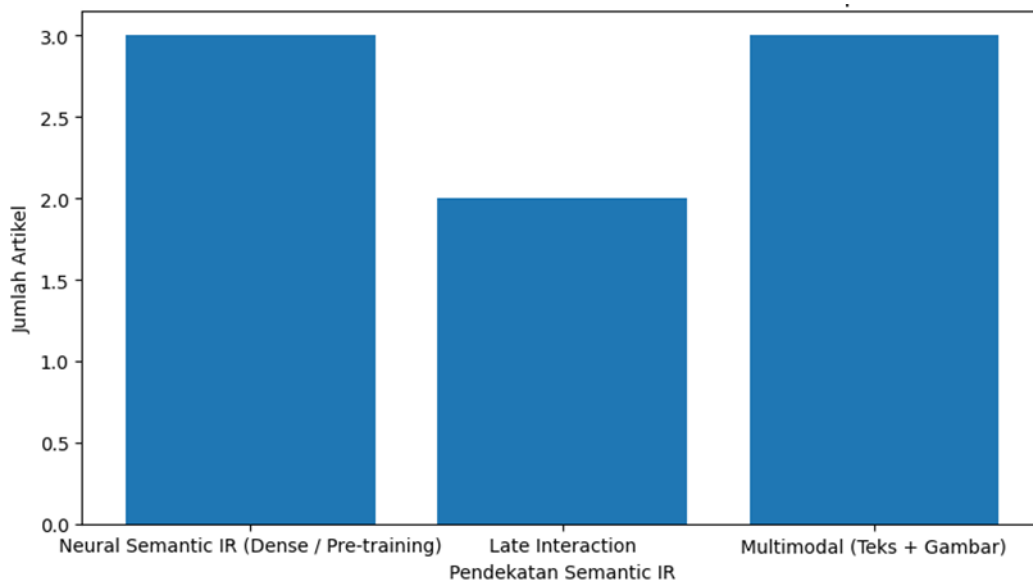
Sementara itu, pendekatan *content-relation graph* muncul dalam jumlah yang lebih terbatas, namun tetap menunjukkan kontribusi penting dalam memperkaya representasi semantik berbasis relasi antar konten. Secara keseluruhan, grafik ini menegaskan bahwa pendekatan *graph-based semantic IR* berperan signifikan dalam meningkatkan pemahaman konteks dan relasi semantik, meskipun memiliki tantangan berupa kompleksitas pemodelan dan kebutuhan data yang besar. Temuan ini memperkuat posisi metode berbasis graf sebagai pendekatan lanjutan dalam pengembangan *semantic IR* untuk pencarian produk berbasis deskripsi.

3.1.5. *Semantic IR Neural dan Multimodal untuk Pencarian Produk Deskriptif*

Pendekatan *semantic IR* paling mutakhir ditunjukkan dalam artikel Liu dan Lopez Ramos (2025), Hu et al. (2024), Don et al. (2024), Liu et al. (2022), serta Gill et al. (2025). Artikel-artikel ini mengembangkan *semantic IR* berbasis *neural network*, *pre-training*, *cluster language model*, dan *late interaction* untuk memahami deskripsi produk secara lebih mendalam [17], [18], [19], [20], [21].

Pendekatan *late interaction* yang diusulkan oleh Gill et al. (2025) secara khusus menunjukkan keunggulan dalam mempertahankan detail semantik token-level, sehingga sangat efektif untuk pencarian produk berbasis deskripsi panjang. Sementara itu, pendekatan multimodal yang menggabungkan teks dan citra produk terbukti mampu meningkatkan relevansi pencarian pada kategori produk tertentu [17], [18].

Meskipun menawarkan performa terbaik, pendekatan ini memiliki kelemahan berupa kompleksitas komputasi dan kebutuhan data yang besar, sehingga implementasinya lebih sesuai untuk sistem e-commerce berskala besar.



Gbr. 7. Distribusi Artikel dengan Pendekatan Semantic IR Neural dan Multimodal untuk Pencarian Produk Deskriptif

Berdasarkan Gambar 7, terlihat bahwa pendekatan *neural semantic IR* dan *multimodal semantic IR* memiliki proporsi yang paling dominan dalam kelompok ini. Hal ini menunjukkan bahwa penelitian terkini dalam pencarian produk berbasis deskripsi cenderung mengandalkan representasi semantik yang dipelajari secara mendalam melalui model neural, serta pemanfaatan berbagai sumber informasi untuk merepresentasikan produk secara lebih komprehensif. Pendekatan ini dinilai efektif dalam menangani deskripsi produk yang panjang, tidak terstruktur, dan mengandung banyak atribut implisit.

Pendekatan *late interaction* meskipun jumlah artikelnnya lebih sedikit, tetap menunjukkan peran yang signifikan. Metode ini menawarkan keunggulan dalam mempertahankan interaksi semantik pada tingkat token antara kueri dan dokumen, sehingga mampu meningkatkan relevansi pencarian pada skenario kueri deskriptif dan *long-tail*. Secara keseluruhan, grafik ini menegaskan bahwa pendekatan semantic IR berbasis neural dan multimodal merupakan arah utama pengembangan sistem pencarian produk e-commerce berbasis deskripsi, meskipun diiringi oleh tantangan berupa kompleksitas komputasi dan kebutuhan data yang besar.

4. Kesimpulan

Berdasarkan analisis terhadap lima belas artikel yang menjadi objek studi literatur, hasil penelitian menunjukkan bahwa pengembangan *Semantic Information Retrieval System* (Semantic IRS) untuk pencarian produk e-commerce berbasis deskripsi didominasi oleh pendekatan berbasis representasi semantik tingkat lanjut. Sebanyak 33% artikel berfokus pada pendekatan *semantic IR* dan *dense retrieval* sebagai fondasi utama pemetaan kueri dan dokumen produk. Pendekatan *late interaction* dan *multimodal semantic retrieval* masing-masing digunakan pada 27% artikel, yang menandakan meningkatnya perhatian terhadap pemodelan interaksi token-level serta integrasi informasi teks dan visual dalam pencarian produk. Sementara itu, pendekatan *query expansion*, *semantic similarity*, dan personalisasi masih relatif terbatas dengan proporsi 13% artikel, menunjukkan bahwa aspek pemahaman kueri masih berperan sebagai pendukung dibandingkan representasi dan pencocokan semantik.

Ditinjau dari karakteristik metodologi, sebanyak 80% artikel menggunakan pendekatan eksperimental kuantitatif dengan evaluasi kinerja sistem berbasis metrik *Information Retrieval System*, dan 67% artikel mengadopsi model neural berbasis transformer. Pendekatan *dense retrieval* digunakan pada 40% artikel, sedangkan *late interaction* dan *multimodal learning* masing-masing muncul pada 27% artikel, mencerminkan pergeseran riset menuju sistem pencarian yang lebih kontekstual dan adaptif terhadap deskripsi produk yang panjang dan tidak terstruktur. Secara keseluruhan, hasil kuantitatif ini menegaskan bahwa pendekatan *Semantic IR* berbasis model neural, *late interaction*, dan multimodal merupakan arah dominan dan paling relevan dalam pengembangan sistem

pencarian produk e-commerce berbasis deskripsi, meskipun diiringi oleh tantangan pada aspek kompleksitas komputasi dan kebutuhan data berskala besar.

Ucapan Terima Kasih

Penulis mengucapkan terima kasih kepada Fakultas Sains dan Teknologi, Institut Teknologi dan Bisnis Indonesia, atas dukungan akademik dan fasilitas yang diberikan sehingga penelitian ini dapat diselesaikan dengan baik. Ucapan terima kasih juga disampaikan kepada seluruh dosen pembimbing dan pihak terkait yang telah memberikan arahan, masukan, serta dukungan ilmiah selama proses penyusunan penelitian ini. Selain itu, penulis mengapresiasi para peneliti dan penerbit jurnal yang karyanya menjadi rujukan utama dalam studi literatur ini, sehingga memperkaya landasan teoretis dan analisis penelitian.

Daftar Pustaka

- [1] O. Rudiansyah, Ariansyah, R., Nanda, R., Wiranda, "Search Engine Menggunakan Metode Information Retrival," *Jurnal Santi - Sistem Informasi Dan Teknik Informasi*, Vol. 2, No. 1, Pp. 49–55, 2022, Doi: 10.58794/Santi.V2i1.68.
- [2] I. M. A. W. Jayarana, I. G. N. A., Darma, I. G. W., Juliantara, I. W. A., & Putra, "Study Literatur Information Retrieval Model: Teknik Dan Aplikasi," *Jurnal Sutasoma*, Vol. 3, No. 1, Pp. 61–69, 2025, Doi: 10.58878/Sutasoma.V3i2.392.
- [3] W. G. Adam, S. I., & Mokodaser, "Implementasi Sistem Rekomendasi Produk E-Commerce Menggunakan Content-Based Filtering Berbasis Cosine Similarity," *Jurnal Sistem Informasi Dan Teknik Komputer*, Vol. 10, No. 2, Pp. 427–434, 2025, Doi: 10.33795/Jip.V11i4.7398.
- [4] A. Syaifuddin, "Sistem Rekomendasi Produk Berbahasa Indonesia Pada Marketplace Tokopedia Menggunakan Metode Content Base Filtering," *Jurnal Ilmiah Teknologi Informasi Dan Sains*, Vol. 3, No. 1, Pp. 70–74, 2023.
- [5] M. A. Herdiansyah, B. W., & Setiawan, "Implementasi Rag Multimodal Untuk Rekomendasi Teks Dan Gambar Di E-Commerce," *Jurnal Pendidikan, Sains Dan Teknologi*, Vol. 12, No. 3, Pp. 1767–1788, 2025.
- [6] A. Z. Mafazi, L., Akhsani, Z., Fadillah, F., Iskandar, D. M., Akbar, Y., & Hidayat, "Pengembangan Fitur Pencarian Dan Filter Produk Pada Aplikasi E-Commerce Gallery Muslim Berbasis Android," *Jurnal Indonesia : Manajemen Informatika Dan Komunikasi*, Vol. 6, No. 3, Pp. 1789–1795, 2025, Doi: 10.63447/Jimik.V6i3.1587.
- [7] A. Rachmaniar, S. Widayati, And K. Rokoyah, "Sistem Rekomendasi Produk E-Commerce Menggunakan Collaborative Filtering Dan Content-Based Filtering (E-Commerce Product Recommendation System Using Collaborative Filtering And Content-Based Filtering)," *Journal Of Information System, Informatics And Computing Issue Period*, Vol. 9, No. 1, Pp. 40–54, 2025, Doi: 10.52362/Jisicom.V9i1.1904.
- [8] L. C. Amalia, A. A., Setyawati, N., & Munggaran, "Sistematik Literatur Sistem Temu Kembali Informasi Dengan Vector Space Model Dan Depth First Search," *Jurnal Ilmiah Komputasi*, Vol. 20, No. 4, Pp. 541–548, 2021, Doi: <https://doi.org/10.32409/jiksti.20.4.2793>.
- [9] T. Ramadhani, S. Nabilah, A. Abimayu, And T. Loi, "Pengembangan Sistem Rekomendasi Produk E-Commerce Menggunakan Algoritma Collaborative Filtering," *Riggs: Journal Of Artificial Intelligence And Digital Business*, Vol. 4, No. 2, Pp. 4848–4854, 2025, Doi: 10.31004/Riggs.V4i2.1349.
- [10] P. Sapanji, R. A. E. V. T., Hamdani, D., & Harahap, "Sentiment Analysis Of The Top 5 E-Commerce Platforms In Indonesia Using Text Mining And Natural Language Processing (Nlp)," *Journal Of Applied Informatics And Computing*, Vol. 7, No. 2, Pp. 202–211, 2023, Doi: 10.30871/Jaic.V7i2.6517.
- [11] F. Kalimantan, M. A. N., & Agustin, "Implentasi Information Retrieval System Menggunakan Metode Latent Semantic," *Jurnal Info Digit*, Vol. 1, No. 2, Pp. 744–755, 2023, [Online]. Available: <http://kti.potensi-utama.ac.id/index.php/jid>
- [12] G. He, Y. Lan, J. Jiang, W. X. Zhao, And J. R. Wen, "Improving Multi-Hop Knowledge Base Question Answering By Learning Intermediate Supervision Signals," *Wsdm 2021 - Proceedings Of The 14th Acm International Conference On Web Search And Data Mining*, No. March 2021, Pp. 553–561, 2021, Doi: 10.1145/3437963.3441753.
- [13] L. Fan Et Al., *Modeling User Behavior With Graph Convolution For Personalized Product Search*, Vol. 1, No. 1. Association For Computing Machinery, 2022. Doi: 10.1145/3485447.3511949.
- [14] G. Xv Et Al., "E-Commerce Search Via Content Collaborative Graph Neural Network," *Proceedings Of The Acm Sigkdd International Conference On Knowledge Discovery And Data Mining*, Pp. 2885–2897, 2023, Doi: 10.1145/3580305.3599320.
- [15] Z. Zhou Et Al., *Semantic-Enhanced Modality-Asymmetric Retrieval For Online E-Commerce Search*, Vol. 1, No. 1. Arxiv, 2023. Doi: 10.1145/3539618.3591863.
- [16] L. Massai, "Evaluation Of Semantic Relations Impact In Query Expansion-Based Retrieval Systems," *Knowl. Based. Syst.*, Vol. 283, No. June 2023, P. 111183, 2024, Doi: 10.1016/J.Knosys.2023.111183.
- [17] D. Liu And E. Lopez Ramos, *Multimodal Semantic Retrieval For Product Search*, Vol. 1, No. 1. Association For Computing Machinery, 2025. Doi: 10.1145/3701716.3717567.
- [18] Z. Hu, S. Li, M. Du, A. Dhua, And D. Gray, "Multimodal Learning With Online Text Cleaning For E-Commerce Product Search 2 Amazon Visual Shopping 3 Amazon Visual Shopping 4 Amazon Visual Shopping 5 Amazon Visual Shopping," 2024.
- [19] D. R. Don, Y. Xie, L. Yu, S. Hughes, And Y. Zhu, "Cluster Language Model For Improved E-Commerce Retrieval And Ranking: Leveraging Query Similarity And Fine-Tuning For Personalized Results," *7th Workshop On E-Commerce And Nlp, Ecnlp 2024 At Lrec-Coling 2024 - Workshop Proceedings*, No. Ecnlp 7, Pp. 145–153, 2024.

- [20] Z. Liu, W. Zhang, Y. Chen, W. Sun, M. Du, And B. Schroeder, "Towards Generalizable Semantic Product Search By Text Similarity Pre-Training On Search Click Logs," *Ecnlp 2022 - 5th Workshop On E-Commerce And Nlp, Proceedings Of The Workshop*, No. Ecnlp 5, Pp. 224–233, 2022, Doi: 10.18653/V1/2022.Ecnlp-1.26.
- [21] A. S. Gill, S. Patel, P. Varga, P. Miller, And S. Athanasiadis, "From Keywords To Concepts: A Late Interaction Approach To Semantic Product Search On Ikea.Com," *Sigir 2025 - Proceedings Of The 48th International Acm Sigir Conference On Research And Development In Information Retrieval*, No. July 2025, Pp. 4280–4283, 2025, Doi: 10.1145/3726302.3731948.