

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pertumbuhan publikasi ilmiah nasional membuat kebutuhan temu-balik informasi yang akurat dan cepat di portal SINTA kian mendesak, karena SINTA berfungsi sebagai sistem informasi riset berbasis web yang menyediakan akses ke publikasi, metrik, dan pemeringkatan jurnal Indonesia; namun, variasi istilah lintas disiplin dan sinonimi sering menyebabkan hasil penelusuran kurang relevan sehingga diperlukan pendekatan yang lebih semantik. (Kemdikbudristek/SINTA, 2025).

Dalam disiplin Information Retrieval (IR) modern, evaluasi kualitas pencarian mensyaratkan metrik ketat—seperti precision, recall, MAP, dan nDCG—serta rancangan eksperimen yang mempertimbangkan karakter koleksi, sehingga strategi peningkatan relevansi harus diuji secara sistematis, bukan sekadar berbasis kecocokan kata kunci. (Sakai, 2021).

Salah satu pendekatan untuk memodelkan “benang merah” makna dalam teks adalah lexical chain, yakni rangkaian keterkaitan leksikal (misalnya sinonimi dan hiponimi) yang merepresentasikan kohesi topikal; studi mutakhir menunjukkan fitur-fitur lexical chain memberi informasi tambahan di luar fitur leksikal standar untuk memahami kesulitan/kohesi teks—indikasi kuat bahwa ia relevan bagi peningkatan relevansi pencarian ilmiah. (Mukherjee, Mitra, & Bhattacharya, 2018).

Implementasi lexical chain memerlukan sumber daya leksikal andal; English WordNet 2019 menghadirkan wordnet berbahasa Inggris dengan model pengembangan open-source yang memperbaiki banyak kekeliruan versi sebelumnya, sehingga lebih siap diintegrasikan ke pipeline NLP modern. (McCrae, Rademaker, Rudnicka, & Bond, 2019).

Pada konteks Indonesia/Melayu, WordNet Bahasa dan sumber leksikal terkait terbukti bermanfaat untuk tugas kemiripan semantik, yang implikasinya adalah jembatan istilah kueri–dokumen dapat dibangun lebih baik ketika istilah Indonesia dan Inggris bercampur dalam artikel SINTA. (Zakaria, Mustapha, & Omar, 2020).

Dari sisi strategi peningkatan relevansi, query expansion (QE)—termasuk yang berbasis WordNet/KB dan hibrida dengan embeddings—terbukti meningkatkan kinerja pencarian pada berbagai koleksi jika seleksi istilahnya dirancang hati-hati dan dievaluasi dengan protokol yang tepat. (Azad & Deepak, 2019).

Akhirnya, praktik perankingan modern juga memanfaatkan kemajuan pemodelan semantik untuk re-ranking hasil; dalam kerangka sistem SINTA, integrasi lexical chain (untuk kohesi topikal) bersama strategi QE yang terkontrol berpotensi meningkatkan ketepatan temu-balik, dan rancangan ini sejalan dengan arah mutakhir perankingan teks. (Lin, Nogueira, & Yates, 2021).

Berdasarkan hasil penelitian dari jurnal relevan, Asrianda, Safwandi, & Raihan (2021) menunjukkan bahwa penggunaan metode *Importance Performance Analysis (IPA)* mampu mengukur kepuasan pengguna dengan baik pada konteks wisata, di mana hasilnya mencapai tingkat kesesuaian sebesar 96% yang berarti pelayanan dan fasilitas yang diberikan sesuai harapan pengunjung. Di sisi lain, Suwanda, Anshari, & Rizal (2022) menekankan pentingnya sistem berbasis web dalam pengelolaan data, khususnya dalam konteks manajemen data alumni. Implementasi *framework CodeIgniter* terbukti dapat menghasilkan sistem yang terstruktur, terkomputerisasi, dan mampu memudahkan proses pencarian serta pelacakan data alumni secara efektif.

Keterkaitan dengan penelitian Putri terlihat pada fokus yang sama dalam penggunaan teknologi untuk meningkatkan kualitas layanan informasi. Baik penelitian berbasis *IPA* pada sektor wisata maupun sistem manajemen alumni berbasis *CodeIgniter* sama-sama menekankan pentingnya efektivitas sistem dalam meningkatkan kepuasan pengguna dan relevansi data. Hal ini selaras dengan upaya pengembangan sistem pencarian jurnal berbasis algoritma tertentu untuk meningkatkan kualitas temu balik informasi akademik (Asrianda et al., 2021; Suwanda et al., 2022).

1.2 Rumusan Masalah

Dimana berikut adalah rumusan masalah yang dijadikan dasar dalam penelitian ini yaitu :

1. Bagaimana mengimplementasikan algoritma *lexical chain* untuk aplikasi pencarian jurnal sinta berbasis web
2. Bagaimana hasil dari implementasi algoritma *lexical chain* untuk aplikasi pencarian jurnal sinta berbasis web

1.3 Tujuan Penelitian

Dimana berikut adalah tujuan penelitian untuk permasalahan diatas dimana sebagai berikut :

1. Mengimplementasikan algoritma *lexical chain* pada aplikasi pencarian jurnal SINTA berbasis web.
2. Mengevaluasi hasil implementasi algoritma *lexical chain* untuk mengetahui efektivitasnya dalam meningkatkan relevansi hasil pencarian jurnal SINTA.

1.4 Manfaat Penelitian

Berikut manfaat penelitian yang diharapkan dalam bisa tercapai setelah proses penelitian ini yaitu :

1. Memberikan kontribusi terhadap pengembangan sistem temu-balik informasi (*information retrieval*) berbasis semantik, khususnya melalui penerapan algoritma *lexical chain* pada pencarian jurnal.
2. Menjadi referensi akademik bagi peneliti berikutnya yang ingin mengembangkan sistem pencarian cerdas di lingkungan publikasi ilmiah Indonesia, terutama pada portal SINTA.
3. Menambah pengalaman dalam penerapan teori dan algoritma *Natural Language Processing (NLP)* pada kasus nyata, sehingga meningkatkan kemampuan teknis dan analitis.
4. Menjadi sarana untuk melatih keterampilan penelitian ilmiah, penulisan akademik, serta penguasaan metode evaluasi sistem berbasis web.

1.5 Ruang Lingkup Dan Batasan Penelitian

Berikut adalah Batasan dari penelitian ini dimana sebagai berikut :

1. Ruang lingkup sistem
Dibatasi pada implementasi algoritma *lexical chain* untuk pencarian jurnal di portal SINTA berbasis web, tanpa mencakup integrasi dengan database lain di luar SINTA.
2. Data uji coba
Hanya menggunakan artikel dan metadata jurnal yang terindeks dalam SINTA, Dimana menggunakan rentang akreditasi SINTA mulai dari SINTA 1 hingga SINTA 6, tidak mencakup sumber publikasi internasional (misalnya Scopus atau WoS).
3. Evaluasi system
Difokuskan pada pengukuran kualitas hasil pencarian menggunakan metrik Information Retrieval (precision, recall, MAP, nDCG), tanpa membahas aspek performa server atau beban sistem secara mendalam.
4. Bahasa penelitian
Dibatasi pada Bahasa Indonesia dan Inggris yang terdapat dalam jurnal SINTA, sehingga pencarian berbasis *lexical chain* hanya diterapkan pada kedua bahasa tersebut.