

BAB I

PENDAHULUAN

1.1. Latar Belakang

Kemampuan dalam melaksanakan penelitian dan mempublikasikan temuannya di jurnal ilmiah bagi dosen dan mahasiswa menjadi keperluan yang penting. Pada perguruan tinggi di Indonesia, setiap dosen diwajibkan untuk menjalankan Tridharma, yang mencakup pengajaran, penelitian, dan pengabdian kepada masyarakat. Bagi para dosen, di antara ketiga aspek tersebut, penelitian sering menjadi tantangan tersendiri. Mereka diharapkan terus menambah keterampilan riset dan mempublikasikan hasilnya di jurnal-jurnal ternama dan terakreditasi, baik nasional maupun internasional yang bereputasi tinggi. Mahasiswa tingkat Magister (S2) dan Doktoral (S3) juga diharuskan memiliki kemampuan untuk mempublikasikan karya ilmiah mereka di jurnal terakreditasi. Bahkan, keterampilan dalam publikasi karya ilmiah kini semakin dianggap penting bagi mahasiswa tingkat Sarjana (S1) [1].

Selain tuntutan tersebut, publikasi jurnal sangat penting karena tidak hanya menjadi bagian dari pencapaian akademik, tetapi juga berkontribusi pada perkembangan ilmu pengetahuan dapat menjadikan kebanggaan tersendiri bagi peneliti. Dengan begitu peneliti berlomba-lomba untuk mempublikasikan hasil karyanya tersebut terutama di jurnal yang terakreditasi. Jurnal yang terakreditasi seperti yang terdaftar di *Scopus* bertebaran banyak di internet, sehingga dapat membuat banyak peneliti sering merasa bingung dan kesulitan dalam menentukan jurnal yang tepat. Hal ini terutama berlaku saat mereka harus mencocokkan topik pada penelitian mereka dengan *scope*, *focus* dan *aim* yang ada pada penjelasan tiap-tiap jurnal. Faktor pertimbangan lainya juga menjadi kendala seperti indeks reputasi jurnal (*Cite Score*), indeksasi *Scopus* (*Quantile*) serta pengaruh journal berdasarkan sitasi (*Impact Factor*) dan biaya publikasi (*APC*). Dengan pemilihan jurnal yang tepat yang sesuai dengan topik apa yang di teliti, peluang diterimanya untuk publikasi pun menjadi tinggi, serta berdampak juga pada durasi waktu publikasi.

Dengan keterbatasan tersebut, diperlukanya sebuah solusi untuk dapat membantu peneliti dalam mendapatkan informasi jurnal yang sesuai. Salah satu solusinya dengan memanfaatkan sistem rekomendasi untuk memberikan saran jurnal yang peneliti harapkan. Dalam konteks ini, sistem rekomendasi memungkinkan peneliti untuk memasukkan abstrak penelitian mereka, yang kemudian diproses menggunakan metode *Content-Based Filtering*. Pendekatan ini bekerja dengan menganalisis kesesuaian konten dari abstrak yang diinput dengan focus, scope, dan aim dari jurnal yang ada, berdasarkan kemiripan fitur-fitur yang relevan dari kedua teks tersebut. Untuk meningkatkan akurasi pencocokan, model *Sentence - Bidirectional Encoder Representations From Transformers (S-BERT)* digunakan untuk menghasilkan representasi vektor dari abstrak dan deskripsi jurnal. *S-BERT* memiliki keunggulan dalam menangkap hubungan semantik antar kalimat, sehingga mampu memahami makna yang lebih dalam dari teks. Setelah itu, teknik *Cosine Similarity* diaplikasikan untuk menghitung sudut antara dua vektor abstrak dan jurnal dalam ruang multidimensi, di mana semakin kecil sudut tersebut, semakin mirip konten kedua teks. Dengan demikian, sistem ini mampu memberikan rekomendasi jurnal yang paling relevan dengan topik penelitian peneliti, serta mempertimbangkan faktor-faktor lain yang penting dalam proses publikasi.

Penelitian sebelumnya yang dilakukan oleh Kim ditahun 2023 [2] menggunakan metode *Content-Based Filtering* berbasis *Sentence-BERT (S-BERT)* untuk merekomendasikan jurnal ilmiah berbahasa Korea. Penelitian ini mampu mengidentifikasi jurnal yang relevan berdasarkan kemiripan abstrak menggunakan vektor embedding *S-BERT* dan menghitung kemiripan semantik menggunakan *Cosine Similarity*. Dataset yang digunakan adalah kumpulan jurnal yang disediakan oleh KISTI dari berbagai artikel berbahasa Korea. Hasil akhir dari penelitian tersebut menunjukkan tingkat akurasi *Top-10 micro* sebesar 76.6%.

Penelitian yang dilakukan oleh Assegaf ditahun 2023 [3] dengan metode *Content-Based Filtering* berbasis *TF-IDF* untuk merekomendasikan jurnal yang sesuai. Dalam studi ini, dataset diambil dari artikel yang terindeks di portal GARUDA. Penelitian ini memanfaatkan *Cosine Similarity* untuk mengukur kesamaan antara abstrak makalah dan artikel jurnal dalam basis data,

menghasilkan rekomendasi jurnal dengan tingkat akurasi yang cukup tinggi, terutama dalam konteks penelitian yang terkait dengan subjek yang lebih luas. Hasil evaluasi menunjukkan bahwa model rekomendasi ini dapat mencapai *Mean Average Precision (MAP)* tertinggi sebesar 95%.

Penelitian yang dilakukan oleh Wang ditahun 2018 [4] memakai metode *Content-Based Filtering* dengan kombinasi seleksi fitur *Chi-Square* dan *Softmax Regression* untuk merekomendasikan jurnal dan konferensi dalam bidang *Computer Science*. Sistem rekomendasi ini memanfaatkan abstrak dari manuskrip yang diberikan untuk menyarankan jurnal yang sesuai, di mana hasil evaluasi menunjukkan bahwa sistem mampu mencapai akurasi 61.37% untuk rekomendasi tiga jurnal teratas (Top-3). Dataset yang digunakan terdiri dari lebih dari 14.000 abstrak artikel yang dikumpulkan menggunakan *web crawler* dari berbagai jurnal dan konferensi di bidang *Computer Science*.

Penelitian sebelumnya oleh Jie Zhu ditahun 2020 [5] menggunakan berbagai model *BERT* yang dioptimalkan dengan *Sentence-BERT* untuk menghasilkan embedding teks ilmiah yang relevan. Dengan *S-BERT*, penelitian ini menerapkan model-model seperti *BERT*, *BioBERT*, dan *SciBERT* untuk mendapatkan vektor 768 dimensi yang dioptimalkan khusus untuk tugas *Semantic Textual Similarity (STS)*, yang berguna dalam mencocokkan abstrak penelitian dengan fokus jurnal. Hasil penelitian ini menunjukkan bahwa metode vektorisasi *S-BERT* memungkinkan pencocokan yang lebih mendalam dan akurat, karena mampu menangkap kesamaan semantik antar kalimat dengan baik, menjadikan *S-BERT* unggul dalam menghasilkan rekomendasi jurnal yang relevan.

Berdasarkan sejumlah penelitian sebelumnya, ditemukan bahwa penggunaan dataset dalam sistem rekomendasi jurnal masih kurang luas dan terbatas. Sebagai contoh, penelitian yang dilakukan oleh Kim et al. tahun 2023 [2] hanya menggunakan data dari jurnal berbahasa Korea dan penelitian Assegaf et al. [3] menggunakan dataset dari portal GARUDA, yang terbatas pada artikel lokal, sehingga belum mencakup cakupan internasional. Selain itu, penelitian tersebut masih kurang memperhatikan faktor penting lainnya, seperti biaya publikasi, skor indeksisasi (*CiteScore*) dan *Quantile*, seperti yang diidentifikasi dalam studi Wang et al. tahun 2018 [4]. Dari sisi metode, hampir semuanya dari penelitian

sebelumnya menggunakan pendekatan *Content-Based Filtering*, yang terbatas pada input pengguna tanpa memperhitungkan perilaku atau preferensi pengguna lain. Dari sini metode *Content-Based Filtering* dianggap lebih baik dalam mencocokkan konten berdasarkan abstrak dalam menawarkan rekomendasi yang lebih kontekstual atau personal. Pemilihan *Sentence-BERT* sebagai model juga didasarkan pada kelebihan yang diidentifikasi dalam penelitian terdahulu, di mana *Sentence-BERT* terbukti lebih baik dalam menghasilkan representasi vektor yang memperhitungkan *Semantik Text Similarity* (STS) secara lebih mendalam dibandingkan terutama dengan model *BERT* lain [5]. Dibandingkan model *BERT* lainnya, *Sentence-BERT* menghasilkan *embedding* vektor dengan dimensi 768, yang cukup tinggi untuk menangkap informasi semantik secara detail namun tetap cepat dan tidak membebani sistem. Dimensi ini memberikan keseimbangan antara kedalaman analisis semantik dan kecepatan komputasi, sehingga memungkinkan sistem untuk menjalankan perhitungan kemiripan antar teks dalam waktu yang relatif singkat. Teknik *Cosine Similarity* juga mampu memberikan hasil yang cukup baik dalam mengukur kemiripan antara abstrak penelitian dan jurnal. Dengan demikian, penelitian ini diharapkan dapat menjadi perbaikan dengan memperluas cakupan dataset melalui penggunaan data jurnal internasional yang terindeks *Scopus*, serta pendekatan dengan *Content-Based Filtering* dan model *Sentence-BERT* yang dirasa sesuai untuk memberikan rekomendasi yang lebih relevan dan bermanfaat bagi peneliti serta mempertimbangkan faktor tambahan seperti biaya publikasi dan reputasi jurnal.

1.2. Perumusan Masalah

1.2.1. Permasalahan

Kondisi dimana peneliti yang sering kebingungan dalam melakukan publikasi jurnal internasional terutama yang terakreditasi *Scopus* karena peneliti harus mencocokkan *focus*, *scope* dan *aim* dari banyaknya jurnal yang tidak mudah untuk dilakukan secara manual. Lalu, pada penelitian sebelumnya telah ada topik yang membahas mengenai rekomendasi jurnal, namun dengan dataset jurnal yang belum mengambil dari indeksasi *Scopus* dan kurang memperhatikan terkait faktor-faktor lain seperti indeks reputasi jurnal (*Cite Score*), indeksasi *Scopus* (*Quantile*), pengaruh jurnal berdasarkan sitasi (*Impact Factor*) dan biaya

publikasi (*APC*) yang dapat juga mempersulit dalam menyesuaikan anggaran dan kebutuhan reputasi.

1.2.2. Metode Usulan

Untuk mengatasi masalah tersebut serta telah melakukan studi literatur dengan hal terkait dan mengumpulkan serta melengkapi dataset jurnal dari pangkalan data Scopus, penelitian ini mengajukan metode *Content-Based Filtering* dengan model *Sentence-BERT* dan teknik pencocokan *Cosine Similarity* untuk membuat rekomendasi jurnal dibidang computer science dari abstrak paper peneliti yang dapat diurutkan juga berdasarkan indeks reputasi (*Cite Score*), indeksasi Scopus (*Quantile*), pengaruh jurnal berdasarkan sitasi (*Impact Factor*) dan biaya publikasi (*APC*).

1.2.3. Pertanyaan Penelitian

Dengan pertimbangan permasalahan dan pendekatan yang sudah dijelaskan, muncul beberapa pertanyaan pada penelitian sebagai berikut

1. Bagaimana kesesuaian rekomendasi jurnal yang dihasilkan oleh metode *Content-Based Filtering* menggunakan model *Sentence-BERT* dibandingkan dengan representasi teks *TF-IDF*?
2. Bagaimana pengaruh *sorting* berdasarkan faktor tambahan seperti *Cite Score* , *Quantile*, *Impact Factor* dan *APC* terhadap relevansi dan preferensi pengguna dalam hasil rekomendasi jurnal?

1.3. Tujuan

Penelitian ini memiliki tujuan untuk mengembangkan sebuah sistem rekomendasi yang bertujuan dalam membantu peneliti untuk menemukan jurnal *Scopus* di bidang *Computer Science* yang sesuai dengan topik penelitian dan *preferensi* mereka. Sistem ini memanfaatkan metode *Content-Based Filtering* dengan model *Sentence-BERT (S-BERT)* untuk mencocokkan abstrak penelitian dengan *focus*, *scope* dan *aim* jurnal, serta menggunakan *Cosine Similarity* untuk mengukur tingkat kesamaan. Selain itu, sistem ini akan mempertimbangkan pengurutan faktor seperti *Cite Score*, *Quantile*, *Impact Factor* dan *APC* untuk memberikan rekomendasi yang lebih relevan dan mendukung proses publikasi secara cepat.

1.4. Manfaat

Penelitian ini diharapkan memberikan manfaat dimana sebagai solusi bagi peneliti, baik dosen maupun mahasiswa, yang sering mengalami kesulitan dalam memilih jurnal yang sesuai untuk publikasi karya ilmiahnya. Dengan adanya sistem rekomendasi berbasis *Content-Based Filtering* yang dilengkapi dengan *S-BERT* dan *Cosine Similarity*, peneliti dapat dengan mudah menemukan jurnal yang relevan dengan abstrak penelitian mereka. Selain itu, penelitian ini memberikan nilai tambah dengan menyajikan informasi penting seperti *CiteScore*, *Quantile* dan *Impact Factor* serta *APC*, sehingga peneliti dapat mempertimbangkan aspek-aspek lain dalam memilih jurnal yang tepat. Dengan demikian, sistem ini dapat mempercepat proses publikasi dan meningkatkan peluang diterimanya karya ilmiah di jurnal terakreditasi, berkontribusi pada peningkatan produktivitas akademik serta pengembangan ilmu pengetahuan.

1.5. Batasan Penelitian

Batasan masalah dari penelitian ini adalah :

1. Hasil rekomendasi hanya mencakup jurnal-jurnal di *Scopus*
2. Bidang jurnal dari rekomendasi hanya di bidang *Computer Science*
3. Dataset untuk rekomendasi di ambil pada bulan Oktober tahun 2024 dan tidak diperbarui otomatis jika ada jurnal baru yang masuk ke *Scopus*

1.6. Sistematika Penulisan

BAB 1 PENDAHULUAN

Dalam bagian ini akan dibahas mengenai sub bab diantaranya latar belakang, perumusan masalah, tujuan, dan manfaat, batasan masalah, pertanyaan penelitian, dan sistematika penelitian.

BAB 2 TINJAUAN PUSTAKA

Pada bagian ini disajikan dasar teori dan metode yang digunakan sebagai dasar dan alat untuk mengatasi masalah serta uraian beberapa hal yang terkait dalam penelitian.

BAB 3 METODE PENELITIAN

Pada bagian ini berisi langkah-langkah yang akan dilakukan pada penelitian seperti tahapan penelitian, variabel penelitian, teknik pengambilan data, metode

yang akan mengolah data dengan berisi langkah-langkahnya, dan flowchart penelitian.

BAB 4 HASIL DAN PEMBAHASAN

Dibagian ini dijelaskan mengenai hasil dari penerapan pemodelan hingga ke sistem yang telah dikembangkan, termasuk pengujian dan evaluasi yang telah dilakukan selama penelitian.

BAB 5 PENUTUP

Bab ini menjelaskan mengenai hasil kesimpulan yang diperoleh dan memberikan saran untuk pengembangan penelitian lebih lanjut.