

BAB I

PENDAHULUAN

1.1. Latar Belakang

Tugas utama seorang dosen adalah melaksanakan Tri Dharma Perguruan Tinggi, yaitu pendidikan dan pengajaran, penelitian, serta pengabdian kepada masyarakat. Pelaksanaan Tri Dharma ini merupakan kewajiban yang harus dipenuhi untuk mencapai visi dan misi perguruan tinggi, dan evaluasinya dilakukan setiap semester melalui pengisian Beban Kinerja Dosen (BKD). Pemerintah, melalui Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi, telah menetapkan kebijakan yang mendorong perguruan tinggi untuk menciptakan dosen dan mahasiswa yang inovatif dalam penelitian dan pengabdian kepada masyarakat, salah satunya melalui program Merdeka Belajar Kampus Merdeka (MBKM). Program ini memberikan peluang bagi dosen dan mahasiswa untuk mengembangkan penelitian ke area yang lebih luas. Bagi dosen, kewajiban untuk mempublikasikan karya ilmiah merupakan bagian penting dalam meningkatkan kualitas akademik dan profesionalisme. Salah satu luaran dari penelitian ini adalah publikasi hasil penelitian pada jurnal bereputasi internasional, seperti jurnal yang terindeks Scopus. Dosen diharapkan terus meningkatkan keterampilan riset dan mempublikasikan hasilnya di jurnal-jurnal ternama yang diakui secara nasional maupun internasional untuk mendukung kemajuan ilmu pengetahuan dan reputasi akademik[1].

Jurnal terakreditasi yang terdaftar di Scopus tersebar luas di internet, sehingga sering kali membuat peneliti merasa bingung dan kesulitan dalam menentukan jurnal yang tepat[2]. Kendala ini juga dialami oleh dosen dan mahasiswa yang ingin mempublikasikan karya ilmiahnya. Salah satu faktor yang memengaruhi publikasi adalah insentif, yang berperan penting dalam pembiayaan proses publikasi. Selain itu, dosen sering kali belum memiliki strategi atau referensi jurnal yang sesuai dengan bidang dan keilmuannya[3]. Hambatan lainnya adalah kebutuhan untuk mencocokkan abstrak dan kata kunci penelitian dengan ruang lingkup dan fokus (*focus and scope*) yang dijelaskan oleh setiap jurnal. Faktor tambahan seperti indeks reputasi jurnal (*cite score*), indeksasi Scopus (*quantile*), dan biaya publikasi (*APC*)

juga menjadi pertimbangan penting. Dengan memilih jurnal yang tepat sesuai dengan abstrak dan kata kunci karya ilmiah, peluang diterimanya karya ilmiah untuk publikasi menjadi lebih tinggi, yang pada akhirnya memengaruhi durasi waktu publikasi secara keseluruhan.

Beberapa penelitian telah mengembangkan sistem rekomendasi, seperti rekomendasi pemilihan konsentrasi penjurusan, pemilihan buku, dan pemilihan film. Namun, hingga saat ini, belum terdapat sistem rekomendasi yang khusus untuk pemilihan jurnal terakreditasi internasional, yang menjadi salah satu tantangan dalam proses publikasi ilmiah, terutama dalam memilih jurnal yang tepat untuk karya penelitian. Sistem rekomendasi memiliki peran penting dalam membantu pengambilan keputusan, termasuk pemilihan jurnal yang sesuai dengan karya ilmiah yang ditulis oleh peneliti.[1]. Keterbatasan tersebut mendorong perlunya solusi untuk membantu peneliti dalam menemukan jurnal yang relevan dengan topik penelitian mereka. Salah satu solusi yang dapat diterapkan adalah sistem rekomendasi berbasis metode *content-based filtering*, yang memungkinkan peneliti untuk memasukkan abstrak dan kata kunci penelitian mereka. Metode ini menganalisis kesesuaian antara abstrak dan kata kunci karya ilmiah dengan ruang lingkup dan fokus (*focus and scope*) jurnal, berdasarkan kemiripan fitur-fitur relevan dari kedua teks. Untuk meningkatkan akurasi pencocokan, digunakan model *TF* (*Term Frequency*) dan *IDF* (*Inverse Document Frequency*), yang dapat menghasilkan representasi vektor dari abstrak dan kata kunci jurnal dan juga dengan ruang lingkup dan fokus (*focus and scope*) jurnal. Kelebihan TF-IDF terletak pada kemampuannya untuk mengidentifikasi kata-kata signifikan dalam dokumen, sambil mengurangi pengaruh kata-kata umum yang kurang informatif. Selanjutnya, teknik *Singular Value Decomposition* (SVD) diterapkan untuk ekstraksi fitur dan pencarian informasi, yang memungkinkan sistem mengurangi dimensi data teks yang besar dan mempertahankan informasi semantik yang relevan, seperti hubungan antara kata-kata yang sinonim atau berkaitan secara semantik. Terakhir, teknik *cosine similarity* digunakan untuk menghitung sudut antara dua vektor abstrak dan kata kunci jurnal dengan *focus and scope* jurnal dalam ruang multidimensi, di mana semakin kecil sudutnya, semakin mirip konten kedua teks tersebut[4]. Dengan demikian, sistem ini dapat memberikan rekomendasi

jurnal yang paling relevan dengan topik penelitian peneliti dan mempertimbangkan faktor-faktor penting dalam proses publikasi.

Penelitian sebelumnya yang dilakukan oleh Badieah et al. pada tahun 2023 untuk mengembangkan model rekomendasi tempat publikasi berbasis konten untuk membantu peneliti memilih jurnal yang tepat menggunakan artikel yang diindeks oleh GARUDA. Model ini memanfaatkan data dari portal GARUDA dan melalui tahapan seperti identifikasi bahasa, pemrosesan data, dan representasi vektor menggunakan TF-IDF untuk menghitung kesamaan artikel dengan jurnal menggunakan *cosine similarity*. Hasil eksperimen menunjukkan bahwa kombinasi judul dan abstrak sebagai input menghasilkan hasil terbaik dalam pemeringkatan jurnal, dengan presisi mencapai 95%. Model ini memberikan rekomendasi jurnal yang relevan berdasarkan kesamaan konten, dengan evaluasi menggunakan atribut nama jurnal dan bidang penelitian. Meskipun model ini efektif, penelitian ini juga mengidentifikasi bahwa pengurangan dimensi melalui teknik ekstraksi kata kunci dapat meningkatkan kinerja model lebih lanjut[5].

Penelitian yang dilakukan oleh Kadhim et al. pada tahun 2017 menggunakan metode TF-IDF yang dikombinasikan dengan teknik *Singular Value Decomposition* (SVD) untuk ekstraksi fitur pada teks di Twitter. Dalam penelitian ini, data yang digunakan terdiri dari 2,196 tweet yang dikelompokkan dalam 12 kategori umum, seperti politik, kesehatan, dan musik. Penelitian ini mengembangkan metode TF-IDF yang dimodifikasi dengan logaritma untuk mengatasi masalah terkait frekuensi kata yang tidak representatif, dan menggunakan SVD untuk mereduksi dimensi data teks. Hasil eksperimen menunjukkan bahwa kombinasi Log (TF-IDF) dan SVD meningkatkan kinerja dalam pengklasifikasian teks, dengan nilai *F1-measure* yang lebih tinggi dibandingkan dengan metode TF-IDF tradisional. Simulasi juga menunjukkan bahwa SVD dapat mengurangi dimensi data tanpa kehilangan fitur penting, meningkatkan efisiensi dalam pemrosesan data berukuran besar. Model yang diusulkan menunjukkan kinerja superior dalam hal akurasi klasifikasi teks, terutama ketika dimensi data ditingkatkan hingga 200 fitur[6].

Penelitian sebelumnya yang dilakukan oleh Darso, Imam Much Ibnu Subroto, dan Sri Arttini Dwi Prasetyowati pada tahun 2017 mengembangkan sistem

rekomendasi publikasi berbasis pengukuran *cosine similarity* untuk portal Indonesian *Publication Index (IPI)* Garuda yang memiliki lebih dari 4000 jurnal Indonesia. Penelitian ini bertujuan untuk mempermudah pengguna dalam menemukan artikel yang relevan melalui perhitungan kesamaan antar dokumen menggunakan pendekatan *cosine similarity*. Proses *preprocessing* data meliputi pembersihan teks, tokenisasi, penghapusan kata-kata stop, dan *stemming*, dengan normalisasi TF-IDF sebagai vektor pembobot untuk perhitungan *cosine similarity*. Hasil eksperimen menunjukkan bahwa sistem ini memiliki *recall* sebesar 0,96 dan *precision* 0,66, yang menunjukkan keberhasilan dalam memberikan rekomendasi yang relevan untuk pengguna[7].

Dari berbagai penelitian sebelumnya, ditemukan bahwa penggunaan dataset dalam sistem rekomendasi jurnal masih terbatas. Penelitian yang dilakukan oleh Badieah et al. pada tahun 2023[5] dan oleh Darso, Imam Much Ibnu Subroto, serta Sri Arttini Dwi Prasetyowati pada tahun 2021[7] hanya menggunakan dataset dari portal GARUDA, yang mencakup artikel lokal dan belum memperhatikan cakupan internasional. Selain itu, penelitian tersebut kurang mempertimbangkan faktor-faktor penting seperti reputasi jurnal (*cite score*), indeksasi Scopus (*quantile*), dan biaya publikasi (*APC*). Dari segi metode, semua penelitian sebelumnya menerapkan pendekatan *content-based filtering* yang hanya berfokus pada input pengguna tanpa mempertimbangkan perilaku atau preferensi pengguna lain. Meskipun metode *content-based filtering* efektif dalam mencocokkan konten berdasarkan judul untuk menawarkan rekomendasi yang lebih kontekstual, pemilihan metode TF-IDF yang dikombinasikan dengan teknik *Singular Value Decomposition* (SVD) untuk ekstraksi fitur teks didasarkan pada keunggulan yang diidentifikasi dalam penelitian terdahulu, di mana kombinasi TF-IDF dan SVD dapat meningkatkan kinerja pengklasifikasian teks dengan nilai *F1-measure* yang lebih tinggi dibandingkan metode TF-IDF tradisional. Simulasi menunjukkan bahwa SVD dapat mengurangi dimensi data tanpa kehilangan fitur penting, sehingga meningkatkan efisiensi pemrosesan data berukuran besar[6]. Teknik *cosine similarity* juga terbukti efektif dalam mengukur kemiripan antara judul penelitian dan jurnal. Oleh karena itu, penelitian ini diharapkan dapat memperbaiki sistem rekomendasi dengan memperluas cakupan dataset melalui penggunaan data jurnal internasional yang

terindeks Scopus serta menerapkan pendekatan *content-based filtering* yang memanfaatkan algoritma TF-IDF, SVD, dan teknik pencocokan *cosine similarity* untuk memberikan rekomendasi jurnal berdasarkan abstrak dan kata kunci jurnal karya ilmiah peneliti, sehingga dapat membantu peneliti dalam memilih jurnal yang optimal dengan mempertimbangkan faktor-faktor penting lainnya.

Dari berbagai penelitian sebelumnya, terungkap bahwa penggunaan dataset dalam sistem rekomendasi jurnal masih terbatas. Sebagai contoh, penelitian yang dilakukan oleh Badieah et al. pada tahun 2023 dan oleh Darso, Imam Much Ibnu Subroto, serta Sri Arttini Dwi Prasetyowati pada tahun 2021 hanya menggunakan dataset dari portal GARUDA, yang mencakup artikel lokal dan belum menjangkau cakupan internasional[5][7]. Selain itu, penelitian tersebut kurang memperhatikan faktor penting lainnya, seperti reputasi jurnal (*cite score*), tingkat penerimaan (*acceptance rate*), indeksasi Scopus (*quantile*), dan biaya publikasi (*APC*). Dari segi metode, semua penelitian sebelumnya menggunakan pendekatan *Content-Based Filtering*, yang hanya mempertimbangkan input pengguna tanpa memperhitungkan perilaku atau preferensi pengguna lain. Oleh karena itu, metode *Content-Based Filtering* dianggap efektif dalam mencocokkan konten berdasarkan judul untuk menawarkan rekomendasi yang lebih kontekstual atau personal[4]. Pemilihan metode TF-IDF yang dikombinasikan dengan teknik *Singular Value Decomposition* (SVD) untuk ekstraksi fitur pada teks didasarkan pada keunggulan yang diidentifikasi dalam penelitian terdahulu, di mana kombinasi TF-IDF dan SVD dapat meningkatkan kinerja pengklasifikasian teks dengan nilai *F1-measure* yang lebih tinggi dibandingkan dengan metode TF-IDF tradisional. Simulasi juga menunjukkan bahwa SVD dapat mengurangi dimensi data tanpa kehilangan fitur penting, sehingga meningkatkan efisiensi dalam pemrosesan data berukuran besar. Teknik *cosine similarity* juga mampu memberikan hasil yang baik dalam mengukur kemiripan antara judul penelitian dan jurnal[7]. Dengan demikian, penelitian ini diharapkan dapat menjadi perbaikan dengan memperluas cakupan dataset melalui penggunaan data jurnal internasional yang terindeks Scopus, serta pendekatan *content-based filtering* yang memanfaatkan algoritma TF-IDF, SVD, dan teknik pencocokan *cosine similarity* untuk memberikan rekomendasi jurnal berdasarkan abstrak dan kata kunci karya ilmiah yang telah disusun oleh peneliti dan juga

memberikan manfaat bagi peneliti dengan mempertimbangkan faktor penting lainnya.

1.2. Rumusan Masalah

1.2.1. Permasalahan

Dalam proses publikasi penelitian, khususnya pada jurnal internasional yang terindeks Scopus, peneliti sering kali menghadapi kebingungan dalam memilih jurnal yang sesuai dengan kebutuhan mereka. Kebingungan ini timbul karena banyaknya pilihan jurnal yang tersedia, yang mengharuskan peneliti untuk melakukan pencarian secara manual, sebuah proses yang memakan waktu dan kurang efisien. Ketidakpastian ini semakin diperburuk ketika peneliti telah menyelesaikan penelitiannya, namun belum menemukan jurnal yang tepat untuk publikasi karya tulis ilmiahnya, yang akhirnya dapat menghambat proses publikasi tersebut. Selain itu, keterbatasan pengetahuan mengenai faktor-faktor penting seperti ruang lingkup dan fokus jurnal (*focus and scope*), reputasi jurnal (*cite score*), indeksasi Scopus (*quantile*), serta biaya publikasi (*APC*), sering kali menjadi hambatan tambahan dalam memastikan kesesuaian antara anggaran, reputasi, dan kebutuhan penelitian. Meskipun sudah ada penelitian sebelumnya yang memberikan rekomendasi jurnal, namun hanya berdasarkan dataset jurnal yang belum terindeks dalam Scopus dan kurang memperhatikan faktor-faktor penting yang telah disebutkan di atas, sehingga solusi yang diberikan belum optimal untuk mengatasi permasalahan yang ada.

1.2.2. Metode Usulan

Berdasarkan permasalahan yang ada, penelitian ini mengusulkan solusi untuk mengatasi hal tersebut dengan melakukan studi literatur yang relevan dan mengumpulkan dataset jurnal dari pangkalan data Scopus di bidang Ilmu Komputer. Penelitian ini mengusulkan penerapan metode *content-based filtering* yang memanfaatkan algoritma TF-IDF, SVD, serta teknik pencocokan *cosine similarity* untuk memberikan rekomendasi jurnal berdasarkan abstrak dan kata kunci dari karya ilmiah peneliti. Rekomendasi jurnal tersebut disusun berdasarkan kesesuaian antara abstrak dan kata kunci karya ilmiah dengan fokus dan ruang lingkup penelitian (*focus and scope*), dengan menambahkan fitur penyesuaian reputasi jurnal (*cite score*), indeksasi Scopus (*quantile*), serta biaya publikasi

(*Article Processing Charge* atau *APC*) yang sesuai dengan preferensi pengguna. Diharapkan solusi ini dapat memudahkan peneliti dalam memilih jurnal yang paling sesuai dan optimal.

1.2.3. Pertanyaan Penelitian

Berdasarkan uraian permasalahan dan metode usulan diatas, dalam penelitian ini merumuskan beberapa pertanyaan sebagai berikut:

1. Seberapa akurat metode *content-based filtering* yang memanfaatkan algoritma TF-IDF, SVD, dan teknik pencocokan *cosine similarity* untuk memberikan rekomendasi jurnal Scopus berdasarkan abstrak dan kata kunci karya ilmiah peneliti/pengguna?
2. Bagaimana pengaruh penerapan SVD dalam meningkatkan skor kemiripan (*similarity score*) dan kesesuaian rekomendasi jurnal Scopus berdasarkan abstrak dan kata kunci karya ilmiah pengguna dengan *Focus and scope* jurnal?

1.3. Tujuan dan Manfaat

Penelitian mengenai implementasi *content-based filtering* untuk rekomendasi jurnal Scopus bidang *computer science* memiliki tujuan dan manfaat. Tujuan dan manfaat tersebut diantaranya sebagai berikut:

1.3.1 Tujuan

Berdasarkan pertanyaan penelitian dan metode yang diusulkan di atas, penelitian ini memiliki tujuan sebagai berikut:

1. Penelitian ini bertujuan untuk mengimplementasikan dan mengevaluasi tingkat akurasi sistem rekomendasi jurnal Scopus di bidang Ilmu Komputer yang berbasis *content-based filtering*, dengan memanfaatkan algoritma TF-IDF, SVD, dan teknik pencocokan *Cosine similarity*. Sistem ini dirancang untuk memberikan rekomendasi jurnal internasional yang terindeks Scopus berdasarkan abstrak dan kata kunci karya ilmiah peneliti, dengan mencocokkan keduanya dengan fokus dan ruang lingkup jurnal yang relevan.
2. Penelitian ini juga bertujuan untuk membandingkan sistem yang hanya menggunakan TF-IDF tanpa penerapan SVD. Hal ini dilakukan untuk mengevaluasi bagaimana penerapan SVD dapat mempengaruhi

peningkatan skor kemiripan dan kesesuaian rekomendasi jurnal berdasarkan abstrak dan kata kunci peneliti. Hasil penelitian diharapkan dapat memberikan wawasan terkait pengaruh SVD dalam meningkatkan kualitas rekomendasi serta memberikan solusi yang lebih efisien bagi peneliti dalam memilih jurnal yang tepat untuk publikasi karya ilmiah mereka.

1.3.2 Manfaat

Berdasarkan tujuan penelitian yang telah dijelaskan, penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Penelitian ini diharapkan dapat memberikan solusi bagi para peneliti di lingkungan akademik, baik dosen maupun mahasiswa, yang sering menghadapi kesulitan dalam memilih jurnal yang sesuai setelah memulai proses penulisan karya ilmiah. Dengan adanya sistem rekomendasi berbasis *content-based filtering* yang memanfaatkan algoritma TF-IDF, SVD, dan *cosine similarity*, peneliti dapat dengan mudah menemukan jurnal yang memiliki fokus dan ruang lingkup yang relevan dengan abstrak dan kata kunci karya ilmiah mereka.
2. Selain itu, sistem ini dilengkapi dengan opsi pengurutan atau sortir tambahan berdasarkan atribut lain, seperti reputasi jurnal (*cite score*), indeksasi Scopus (*quantile*), dan biaya publikasi (*APC*), yang memungkinkan peneliti untuk mempertimbangkan berbagai aspek dalam memilih jurnal yang tepat. Dengan demikian, sistem ini diharapkan dapat mempercepat proses publikasi, meningkatkan efisiensi waktu, serta memberikan kontribusi signifikan dalam mendorong produktivitas akademik dan kemajuan ilmu pengetahuan.

1.4. Batasan Masalah

Pada penelitian ini terdapat batasan masalah agar tidak terjadi kesalahan pada saat penelitian ini dilakukan. Berikut batasan masalah pada penelitian tersebut:

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

3. Metode yang digunakan yaitu *Content Based Filtering* dengan penggabungan algoritma TF-IDF, SVD dan *Cosine similarity*.

1.5. Sistematika Penulisan

Untuk mempermudah dalam memahami skripsi ini, maka dalam penulisannya dibedakan menjadi beberapa bab, yaitu:

1. BAB I PENDAHULUAN

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

2. BAB II 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, serta penelitian terkait.

3. BAB III METODOLOGI 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.

4. BAB IV HASIL DAN PEMBAHASAN

Bagian ini menjelaskan hasil dari penerapan *modeling* hingga sistem yang telah dikembangkan, Termasuk *testing* dan *evaluation* yang dilakukan selama proses penelitian.

5. BAB V PENUTUP

Bab ini memuat simpulan dari hasil penelitian dan menyampaikan saran sebagai bahan pertimbangan untuk pengembangan penelitian selanjutnya.