

# **BAB I**

## **PENDAHULUAN**

Bab ini memberikan penjelasan lengkap tentang latar belakang, rumusan masalah, tujuan, manfaat, dan batasan penelitian yang dilakukan. Penjelasan dalam bab ini bertujuan untuk memberikan pemahaman awal yang menyeluruh tentang urgensi perancangan sistem rekomendasi topik skripsi, serta ruang lingkup dan pendekatan utama yang diterapkan dalam penelitian ini. Dengan memahami konteks dan dasar dari penelitian ini, diharapkan pembaca akan memiliki gambaran yang jelas sebelum memasuki pembahasan teknis di bab-bab selanjutnya.

### **1.1. Latar Belakang**

Pemilihan topik skripsi merupakan salah satu tahapan paling penting dalam penyusunan tugas akhir, namun sering kali menjadi hambatan besar bagi mahasiswa. Banyak mahasiswa mengalami kebingungan dalam menentukan topik yang sesuai dengan minat, kompetensi, dan bidang keilmuan yang ditekuni. Menurut suatu penelitian, sekitar 60% mahasiswa mengakui kesulitan dalam memilih tema penelitian, yang pada akhirnya berujung pada keterlambatan dalam menyelesaikan studi [1]. Kesulitan ini bukan fenomena individual, tetapi merupakan masalah umum yang dialami mahasiswa di berbagai perguruan tinggi. Menurut studi yang membahas tantangan mahasiswa dalam menentukan judul skripsi, kurangnya pemahaman terhadap bidang studi dan ketidakmampuan mengidentifikasi topik yang relevan menjadi penyebab utama kebingungan mahasiswa dalam memilih topik penelitian [2]. Hal ini berdampak langsung pada keterlambatan proses penyusunan skripsi dan memperpanjang masa studi.

Penelitian lain juga menegaskan bahwa kesulitan dalam menentukan topik merupakan hambatan awal yang signifikan dalam proses penulisan skripsi. Banyak mahasiswa mengganti topik di tengah jalan karena merasa topik awal terlalu sulit atau tidak sesuai dengan minat, sehingga proses penelitian menjadi tidak efisien dan memakan waktu lebih lama [3]. Kondisi ini menunjukkan bahwa pemilihan topik bukan hanya persoalan akademik, tetapi juga persoalan psikologis dan manajemen waktu yang memengaruhi kelancaran studi mahasiswa.

*Repository* kampus sebenarnya menyimpan ribuan judul dan abstrak skripsi yang dapat dimanfaatkan sebagai referensi. Namun, sistem pencarian yang tersedia

umumnya masih bersifat statis dan hanya mengandalkan pencocokan kata kunci sederhana. Mahasiswa harus membaca banyak abstrak secara manual untuk menemukan topik yang relevan, yang tentu memakan waktu dan tidak efisien. Minimnya platform digital yang dapat menjembatani minat mahasiswa dengan topik skripsi secara personal memperburuk situasi ini. Artikel akademik mengenai sistem pendukung keputusan untuk pemilihan topik skripsi menegaskan bahwa pemilihan topik yang tepat merupakan langkah krusial, namun mahasiswa sering kesulitan karena keterbatasan informasi dan tidak adanya alat bantu yang memadai [4].

Dalam mengatasi permasalahan tersebut, pendekatan berbasis algoritma seperti *Term Frequency-Inverse Document Frequency (TF-IDF)* dan *Cosine Similarity* dapat memberikan solusi yang lebih tepat [5]. Penerapan kedua algoritma ini dalam sistem rekomendasi berbasis konten terbukti mampu menghasilkan rekomendasi yang akurat dan terarah, sehingga sangat relevan jika diterapkan pada sistem pemilihan topik skripsi yang sesuai dengan minat dan keahlian mahasiswa [6]. *TF-IDF* merupakan metode pembobotan kata dalam dokumen yang digunakan untuk menentukan pentingnya suatu kata terhadap dokumen dalam kumpulan dokumen tertentu. Bersama dengan *Cosine Similarity*, yang mengukur derajat kemiripan antara dua vektor teks, kedua algoritma ini sangat populer dalam sistem rekomendasi berbasis konten *CBF(content-based filtering)*.

Kombinasi *TF-IDF* dan *Cosine Similarity* dalam sistem rekomendasi resep masakan terbukti mampu menghasilkan rekomendasi yang sangat relevan dengan preferensi pengguna [7]. Hasil serupa juga ditemukan pada sistem rekomendasi lowongan pekerjaan, di mana kedua algoritma tersebut berhasil mencapai akurasi rekomendasi yang tinggi [8]. Di bidang akademik, sistem informasi berbasis *CBF* telah terbukti efektif. Penerapan *TF-IDF* dan *Cosine Similarity* dalam sistem rekomendasi buku perpustakaan kampus berdasarkan minat mahasiswa terbukti sangat efektif [9]. Selain itu, penelitian lain mengembangkan bank judul skripsi menggunakan metode *ADDIE* untuk membantu mahasiswa dalam memilih topik penelitian. Namun, sistem tersebut masih bersifat pasif dan belum mengimplementasikan algoritma pencocokan secara otomatis [10].

Beberapa algoritma lain yang sering digunakan dalam pengukuran kesamaan teks adalah *Jaccard Similarity* dan *Euclidean Distance*. *Jaccard Similarity*, yang mengukur kesamaan berdasarkan keberadaan kata dalam dua set dokumen,

memberikan hasil yang relatif sederhana. Namun, karena hanya memperhitungkan keberadaan kata tanpa mempertimbangkan frekuensi kata, *Jaccard* kurang efektif dalam menangani teks dengan variasi frekuensi kata yang signifikan. Hal ini menjadikannya tidak ideal untuk teks yang lebih kompleks, seperti yang ada dalam sistem rekomendasi topik skripsi, di mana frekuensi kata dapat mencerminkan relevansi dan bobot dari setiap istilah [11]. Di sisi lain, *Euclidean Distance* mengukur jarak fisik antar vektor berdasarkan frekuensi kata dalam dokumen. Meskipun memberikan ukuran jarak antara dua vektor, metode ini sangat dipengaruhi oleh panjang dokumen, yang dapat menyebabkan bias pada perbandingan dokumen dengan ukuran yang sangat berbeda. Hal ini membuat *Euclidean Distance* kurang tepat dalam konteks perbandingan dokumen teks yang bervariasi panjangnya [12]. Oleh karena itu, meskipun kedua algoritma ini berguna dalam konteks tertentu, *Cosine Similarity* lebih unggul untuk sistem rekomendasi topik skripsi. *Cosine Similarity* mengukur kesamaan antar dua vektor berdasarkan sudut antara keduanya, yang menjadikannya lebih tahan terhadap perbedaan panjang dokumen dan lebih akurat dalam mengukur kesamaan berdasarkan frekuensi kata.

Metode modern seperti *word embeddings* dan model berbasis transformer seperti *BERT* juga telah digunakan dalam berbagai penelitian untuk memahami konteks kata secara lebih mendalam. Meskipun metode ini menawarkan akurasi yang tinggi, penerapannya membutuhkan sumber daya komputasi yang besar, proses pelatihan yang kompleks, serta infrastruktur yang lebih mahal [13]. Dalam pengembangan sistem rekomendasi skripsi yang ringan, cepat, dan mudah diimplementasikan di lingkungan kampus, TF-IDF dan *Cosine Similarity* tetap menjadi pilihan praktis tanpa mengorbankan kualitas rekomendasi.

Penelitian ini menggabungkan TF-IDF dan *Cosine Similarity* untuk memberikan analisis lebih mendalam terhadap kemiripan konten dan menghasilkan rekomendasi yang lebih relevan dengan minat mahasiswa. Dengan penerapan sistem berbasis *web* ini, mahasiswa dapat memperoleh rekomendasi topik skripsi secara otomatis, cepat, dan sesuai preferensi pengguna. Berdasarkan berbagai penelitian terdahulu, penggabungan metode TF-IDF dan *Cosine Similarity* memiliki potensi besar untuk diterapkan dalam sistem rekomendasi topik skripsi. Penelitian ini bertujuan untuk mengembangkan sistem yang mampu memberikan rekomendasi topik skripsi secara personal berdasarkan minat mahasiswa. Sistem ini diharapkan

dapat menjadi solusi atas kesulitan yang dialami mahasiswa dalam menentukan topik penelitian, serta mendukung proses akademik secara efisien dan cerdas.

## **1.2. Rumusan Masalah**

Berdasarkan permasalahan yang diidentifikasi pada latar belakang, penelitian ini merumuskan beberapa pertanyaan utama yang akan dijawab, yaitu:

1. Bagaimana merancang dan membangun sistem rekomendasi topik skripsi berbasis *web* yang mampu memberikan rekomendasi topik yang sesuai dengan preferensi, minat, dan bidang keahlian mahasiswa?
2. Bagaimana menerapkan algoritma TF-IDF dan *Cosine Similarity* untuk mencocokkan input mahasiswa dengan deskripsi topik skripsi sehingga menghasilkan rekomendasi yang relevan?
3. Bagaimana hasil performa model rekomendasi dalam menghasilkan topik skripsi yang tepat dan sesuai dengan preferensi mahasiswa?

## **1.3. Tujuan Penelitian**

Penelitian ini bertujuan untuk memberikan kontribusi terhadap solusi digital dalam bidang pendidikan, khususnya dalam membantu mahasiswa menentukan topik skripsi yang tepat. Berdasarkan rumusan masalah yang dijelaskan sebelumnya, tujuan utama yang ingin dicapai pada penelitian ini adalah sebagai berikut ini:

1. Merancang dan mengembangkan sistem rekomendasi topik skripsi berbasis *web* yang dapat memberikan saran topik yang sesuai dengan minat dan bidang keahlian mahasiswa secara relevan.
2. Mengimplementasikan algoritma TF-IDF dan *Cosine Similarity* untuk mengukur relevansi antara input preferensi mahasiswa dengan kumpulan topik skripsi yang ada.
3. Menganalisis dan menyajikan performa model rekomendasi dalam menghasilkan topik skripsi yang tepat dengan preferensi mahasiswa.

## **1.4. Batasan Masalah**

Agar penelitian ini memiliki fokus yang jelas dan dapat diselesaikan secara terukur, maka perlu ditetapkan beberapa batasan masalah sebagai berikut:

1. Penelitian ini hanya mencakup pengembangan sistem rekomendasi topik skripsi berbasis *web* yang diperuntukkan bagi mahasiswa Strata-1 (S1) di Fakultas Ilmu Komputer.

2. Dataset yang digunakan dalam penelitian ini diolah dari beberapa *repository* kampus. Data dikumpulkan dari universitas-universitas di Indonesia dengan total  $\pm 4000$  data, Data yang dikumpulkan meliputi *metadata* tiap skripsi yang terkait dengan bidang ilmu di Fakultas Ilmu Komputer.
3. Algoritma yang digunakan untuk proses rekomendasi yaitu TF-IDF dan *Cosine Similarity*, tanpa melakukan perbandingan dengan algoritma lainnya.
4. Evaluasi performa sistem rekomendasi diperoleh melalui *peer evaluation* oleh pengguna akhir menggunakan kuesioner serta divalidasi oleh *expert reviewer*. Selain itu, juga dilakukan pengujian *blackbox* untuk menguji seluruh komponen dalam sistem.

### 1.5. Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan kontribusi yang nyata, baik secara teoritis maupun praktis, khususnya dalam bidang pengembangan teknologi informasi di lingkungan pendidikan tinggi. Adapun manfaat yang dapat diperoleh dari penelitian ini adalah sebagai berikut:

1. Hasil dari pengembangan sistem rekomendasi topik skripsi dengan penerapan algoritma TF-IDF dan *Cosine Similarity* diharapkan dapat memberikan wawasan baru mengenai efektivitas metode CBF dalam mencocokkan minat mahasiswa dengan topik-topik skripsi yang relevan secara personal dan sistematis.
2. Penelitian ini diharapkan dapat memberikan dukungan nyata bagi mahasiswa dan dosen pembimbing. Dengan menghadirkan sistem yang memudahkan mahasiswa dalam menemukan topik skripsi yang sesuai dengan minat dan kemampuan akademik, sistem ini juga membantu dosen dalam memberikan arahan yang lebih terfokus dan tepat sasaran. Hal ini akan mempercepat proses pemilihan topik skripsi, sehingga tahap awal penyusunan skripsi dapat berlangsung lebih cepat dan terarah.
3. Penelitian ini juga diharapkan dapat menjadi referensi serta pijakan awal bagi peneliti atau pengembang sistem rekomendasi lainnya yang ingin mengkaji penerapan algoritma pencocokan teks terkhusus dalam bidang akademik.