

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Pengajuan judul tugas akhir menjadi tahapan awal dalam rangkaian penyelesaian tugas akhir. Dalam proses pengajuan judul biasanya mahasiswa sudah mempertimbangkan judul yang akan dikerjakan dan disertai dengan *abstract* sebagai gambaran umum dosen dalam menentukan lolos tidaknya judul sehingga bisa masuk ke tahapan selanjutnya untuk proposal dan sampai tahap pengerjaan per bab nya.

Pencarian judul yang beragam membuat sebagai mahasiswa terkadang kesulitan dengan pokok bahasan yang akan dibahas. Salah satu kendala bagi mahasiswa dalam pengambilan judul atau topik bahasan adalah kondisi kurang berkembangnya bidang keilmuan tanpa adanya inovasi baru atau hanya sebatas permasalahan yang diganti objek.

Hal ini akan menimbulkan permasalahan yang lebih besar kedepan nya dikarenakan adanya kemiripan setiap penelitian khususnya dalam hal bidang penelitian. Lebih jauh bisa mengakibatkan suatu plagiat dikarenakan adanya banyak kemiripan yang terjadi. Plagiarisme merupakan salah satu perbuatan yang melanggar kode etik dalam dunia penulisan, karena tindakan tersebut telah mengambil karya orang lain sebagai karya sendiri. Plagiarisme merupakan tindakan mengambil sebagian atau seluruh ide seseorang berupa dokumen maupun teks tanpa mencantumkan sumber pengambilan informasi [1].

Proses pendeteksian kemiripan judul yang sudah dilakukan oleh

Universitas Dinamika Bangsa masih sebatas membandingkan judul tugas akhir mahasiswa oleh dosen pembimbing masing - masing. Hal ini berdampak kurang lengkapnya sumber informasi judul yang sudah di ajukan dalam basis data yang tersedia di Universitas Dinamika Bangsa. Setiap dosen akan mempunyai preferensi penilaian yang berbeda - beda terhadap proposal judul yang diajukan, apakah layak diangkat sebagai judul tugas akhir atau tidak tanpa adanya variabel pengecekan tentang kemiripan judul yang sudah pernah diajukan.

Proses pencarian kemiripan judul dapat menjadi salah satu penyebab penolakan proposal judul tugas akhir. Namun begitu, penggunaan metode yang tepat dapat membantu dalam proses pengkategorian kemiripan dokumen. Salah satu metode yang digunakan untuk proses perhitungan kemiripan teks khususnya untuk judul adalah metode *Cosine Similarity*. *Cosine Similarity* adalah perhitungan kesamaan antara dua vektor n dimensi dengan mencari kosinus dari sudut diantara keduanya dan sering digunakan untuk membandingkan dokumen dalam text mining [2]. Penggunaan *cosine similarity* dapat di kombinasikan dengan pembobotan setiap kata dengan metode *Term Frequency Inverse Document Frequency* (TF-IDF). Metode TF-IDF merupakan metode untuk menghitung bobot suatu kata (*term*) terhadap dokumen. Metode ini juga terkenal efisien, mudah dan memiliki hasil yang akurat [3]. Frekuensi kemunculan kata di dalam dokumen yang diberikan menunjukkan seberapa penting kata tersebut di dalam dokumen.

Penelitian ini dilakukan untuk menguji tingkat kemiripan judul tugas akhir beserta dengan *abstract* yang akan diajukan dengan dibandingkan dengan judul

yang sudah ada di Universitas Dinamika Bangsa. Hasil kemiripan dokumen akan dibagi ke dalam tiga tingkatan kemiripan judul yang meliputi kemiripan rendah, sedang dan tinggi. Berdasarkan uraian di atas maka peneliti mengambil judul "Deteksi Kemiripan Dokumen Teks judul tugas akhir dan *Abstract* dengan Metode *Cosine Similarity* dan Pembobotan *Term Frequency-Inverse Document Frequency* (TF-IDF) pada jurusan Sistem Komputer Fakultas Ilmu Komputer Universitas Dinamika Bangsa".

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas dapat dirumuskan suatu permasalahan yaitu: “Bagaimana merancang dan membangun aplikasi yang mampu mendeteksi tingkat kemiripan judul tugas akhir dengan metode *cosine similarity* dan pembobotan TF-IDF berdasarkan judul dan *abstract*?”

1.3 Batasan Masalah

Batasan masalah dalam penulisan ini dimaksudkan agar pembahasan dan penulisan laporan dapat dilakukan secara terarah dan mencapai sasaran antara lain :

1. Aplikasi yang dibangun adalah aplikasi yang mampu deteksi kemiripan teks judul tugas akhir di Universitas Dinamika Bangsa;
2. Data diambil dari periode tahun 2021 sampai dengan 2022 pada jurusan Sistem Komputer;
3. Metode yang digunakan untuk deteksi kemiripan menggunakan *cosine similarity* dengan pembobotan TF-IDF;

4. Hasil dari kemiripan dikategorikan ke dalam 3 macam jenis yaitu berupa kemiripan rendah, kemiripan sedang dan kemiripan tinggi;
5. Tahapan *stemming* menggunakan *library* tambahan yaitu *library* sastrawi.
6. Aplikasi yang dibangun ber basis web dengan bahasa pemrograman PHP dan basis data MySQL.

1.4 Tujuan dan Manfaat Penelitian

Adapun tujuan dari Penelitian ini sebagai berikut :

1. Membuat suatu sistem yang dapat mendeteksi tingkat kemiripan dari judul tugas akhir dan *abstract* yang diajukan dibandingkan dengan basis data judul tugas akhir dan *abstract* yang ada di Universitas Dinamika Bangsa.
2. Menguji tingkat kemiripan judul dan *abstract* tugas akhir dengan membandingkan *text processing* dengan tahapan *stemming* dibandingkan tanpa tahapan *stemming*.

Sedangkan manfaat dari Penelitian ini sebagai berikut :

1. Dapat membantu pihak Universitas Dinamika Bangsa dalam proses pemilihan judul tugas akhir berdasarkan tingkat kemiripan yang ada.
2. Untuk memperkaya ilmu pengetahuan khususnya dalam hal bidang *text mining* dalam hal deteksi kemiripan *text*.

1.5 Sistematika Penulisan

Sistematika di dalam penulisan penelitian ini dapat dipaparkan sebagai berikut :

BAB I PENDAHULUAN

Pada bab ini akan diuraikan mengenai latar belakang penelitian,

rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan.

BAB II LANDASAN TEORI

Pada bab ini akan diuraikan landasan teori-teori yang mendasari pembahasan secara detail terkait serta studi literatur mengenai perbandingan penelitian terkait

BAB III METODE PENELITIAN

Pada bab ini akan dibahas metode penelitian yang akan digunakan mencakup metode yang akan digunakan dan alat untuk kebutuhan penelitian.

BAB IV ANALISIS DAN PERANCANGAN

Pada bab ini membahas tentang analisis dari data yang digunakan meliputi pengolahan data dan rancangan dari tools yang akan dibuat untuk kebutuhan implementasi dan pengujian.

BAB V IMPLEMENTASI DAN PENGUJIAN SISTEM

Pada bab ini membahas tentang tahapan implementasi dan pengujian dari sistem atau tools yang dirancang pada bab sebelumnya.

BAB VI PENUTUP

Pada bab ini membahas kesimpulan dan saran dengan penelitian