

BAB I PENDAHULUAN

1.1 Latar Belakang

Skripsi merupakan salah satu tahapan akhir dalam menyelesaikan studi di perguruan tinggi. Setiap mahasiswa pada jalur skripsi harus menentukan tema dan judul skripsi sesuai dengan bidang ilmu masing-masing mahasiswa. Skripsi harus ditulis dengan orisinal dalam arti tidak plagiat dan tidak memiliki kemiripan yang sangat tinggi antara satu penelitian dengan penelitian lainnya.

Setiap judul skripsi yang diajukan oleh mahasiswa akan di filter oleh ketua Kelompok Bidang Keilmuan (KBK). Di Darmajaya khususnya prodi Teknik Informatika terdapat 3 konsentrasi bidang keilmuan yaitu Artificial Intelligence, Mobile Technology, dan Multimedia. Masing-masing bidang keilmuan memiliki ketua KBK. KBK akan melakukan pengecekan dari setiap judul skripsi yang diajukan oleh mahasiswa dalam rangka menghindari judul identik, kemiripan yang sangat tinggi, dan plagiarisme dalam penelitian skripsi.

Masalah yang sering terjadi yaitu judul yang diajukan oleh mahasiswa seringkali identik, sangat mirip dengan judul yang sudah pernah diajukan sebelumnya, ataupun mahasiswa hanya mengganti judul dengan lokasi penelitian yang berbeda tanpa adanya nilai keterbaruan dari penelitian tersebut. Judul skripsi yang identik, sangat mirip, dan plagiat pasti ditolak oleh KBK. Saat ini terdapat ±2148 judul skripsi pada *spreadsheet* prodi teknik informatika. Sehingga menyulitkan untuk mendeteksi kemiripan pada judul skripsi yang diajukan. Untuk mendeteksi kemiripan pada judul skripsi.

Kombinasi algoritma *TF-IDF* dengan *Fuzzy Matching* merupakan salah satu algoritma yang cocok untuk mendeteksi tingkat kemiripan judul skripsi yang diajukan oleh mahasiswa.

Algoritma *TF-IDF* adalah metode yang digunakan untuk mengukur seberapa penting sebuah kata dalam sebuah dokumen [1], dimana *TF-IDF* secara umum didefinisikan sebagai 2 algoritma yang berbeda. *TF (Term Frequency)* merupakan algoritma yang menghitung seberapa sering sebuah kata muncul dalam sebuah dokumen, sedangkan *IDF (Inverse Document Frequency)* merupakan algoritma yang menghitung seberapa jarang sebuah kata muncul dalam sebuah dokumen. Algoritma kombinasi lainnya yaitu *Fuzzy Matching* merupakan metode pencocokan string yang menggunakan pendekatan terhadap pola dari string yang dicari, dimana algoritma *Fuzzy Matching* bekerja dengan cara yang berbeda dari pencocokan string biasa (*exact matching*) [2], algoritma ini berfokus pada menemukan kecocokan string yang “mirip”. Sehingga apabila kedua algoritma ini dikombinasikan maka akan menghasilkan tingkat akurasi yang cukup tinggi dalam mendeteksi kemiripan judul – judul skripsi mahasiswa.

Berdasarkan uraian diatas, maka penulis mengajukan judul skripsi **“IMPLEMENTASI ALGORITMA *TF-IDF* DAN *FUZZY MATCHING* UNTUK DETEKSI KEMIRIPAN JUDUL SKRIPSI BERBASIS APLIKASI ANDROID. STUDI KASUS: JUDUL SKRIPSI PRODI TEKNIK INFORMATIKA IIB DARMAJAYA”** untuk memudahkan kelompok bidang keilmuan (KBK) melakukan deteksi tingkat kemiripan pada judul – judul yang diajukan yang ada pada jurusan teknik informatika dengan algoritma *TF-IDF* dan *Fuzzy Matching*.

1.2 Identifikasi Masalah

Berdasarkan uraian pada latar belakang, ditemukan informasi mengenai sistem yang akan dibangun. Identifikasi masalah melalui latar belakang diantaranya:

1. Sulitnya bagi mahasiswa untuk mencari kemiripan ide judul mereka dengan judul yang sudah ada.
2. Pengecekan judul – judul mahasiswa yang diajukan masih dilakukan dengan cara manual yang tersimpan dalam bentuk *Microsoft Excel*.

3. Pertimbangan judul-judul yang telah diajukan ke kelompok bidang keilmuan (KBK) menjadi kurang efisien karena dilakukan satu persatu apabila judul yang diajukan sangat banyak, sehingga memungkinkan untuk terjadinya judul yang sama.

1.3 Rumusan Masalah

Perumusan masalah berdasarkan uraian latar belakang menjadi “Bagaimana kelompok bidang keilmuan (KBK) dapat mendeteksi kemiripan judul yang diajukan oleh mahasiswa teknik informatika dengan judul – judul yang sudah ada dan bagaimana mahasiswa dapat menentukan serta mendeteksi adanya kemiripan ide judulnya dengan judul yang sudah ada dengan menerapkan algoritma *TF-IDF* dan *Fuzzy Matching* untuk aplikasi berbasis android?”.

1.4 Batasan Masalah

Setiap penelitian memiliki batasan – batasan masalahnya. Batasan masalah untuk penelitian ini adalah:

1. Ruang lingkup penelitian ini hanya mencakup program studi teknik informatika IIB Darmajaya.
2. Pengguna aplikasi ini adalah kelompok bidang keilmuan (KBK) dan mahasiswa program studi teknik informatika.
3. Kombinasi algoritma *TF-IDF* dan *Fuzzy Matching* memiliki kekurangan dalam memahami konteks judul – judul, data yang belum pernah ada, hubungan antar judul yang kompleks, serta sumber daya yang tinggi untuk melakukan komputasi dengan algoritma ini.
4. Menggunakan Bahasa pemrograman *Python* sebagai *backend computing* dan database *JSON* untuk penyimpanan data judul.

1.5 Tujuan Penelitian

Tujuan dari penelitian ini ialah:

1. Membantu menghindari adanya kemungkinan judul skripsi yang sama antara judul yang diajukan dengan judul yang sudah ada.
2. Meningkatkan efisiensi proses pertimbangan terhadap judul – judul yang diajukan oleh mahasiswa teknik informatika.
3. Membantu pendeteksian kemiripan judul yang diajukan secara efektif.

1.6 Manfaat Penelitian

Beberapa manfaat dari penelitian ini diantaranya ialah:

1. Membantu pengambilan keputusan pertimbangan kelompok bidang keilmuan (KBK) dalam menentukan judul – judul yang diajukan.
2. Membantu memberikan rekomendasi judul serupa yang cukup akurat.
3. Membantu dalam mendeteksi persentase kemiripan dari judul – judul yang diajukan.

1.7 Sistematika Penulisan

Sistematika penulisan yang digunakan dalam penyusunan penelitian ini adalah sebagai berikut:

BAB I PENDAHULUAN

Berisi penjelasan mengenai latar belakang identifikasi masalah, rumusan masalah, ruang lingkup penelitian, tujuan, manfaat dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Berisi dasar teori yang digunakan sebagai acuan dalam pengambilan data, analisis, hasil penelitian dan pengujian.

BAB III METODE PENELITIAN

Dalam bab ini dijelaskan tentang metode pendekatan-pendekatan yang dilakukan untuk mendukung dalam proses penelitian.

DAFTAR PUSTAKA