

BAB I

PENDAHULUAN

1.1 Latar Belakang

Semakin mudahnya pertukaran informasi dari waktu ke waktu membawa dampak yang positif bagi kemajuan teknologi [1, 2]. Saat ini informasi yang paling banyak digunakan berupa teks dan tidak menutup kemungkinan terdapat informasi yang sama atau mirip. Kemiripan teks dapat mengakibatkan dugaan meniru karya seseorang. Pemicu munculnya teks-teks yang sama, salah satunya karena masyarakat saat ini sering melakukan tindakan *copy-paste*. Tanpa disadari tindakan ini cenderung membuat seseorang menjadi malas untuk berpikir dan bekerja. Tindakan ini harus terus menerus ditekan agar tidak menjadi kebiasaan yang dapat menghambat kreativitas seseorang. Oleh karena itu, diperlukan cara yang efektif untuk mengecek kemiripan pada dokumen teks. Salah satunya dengan memanfaatkan Algoritma pencocokan string yang dapat dikembangkan menjadi sebuah aplikasi untuk mendeteksi kemiripan pada dokumen teks. Dengan mengetahui hasil persentase kemiripan dari dokumen yang diuji, hasil tersebut dapat dijadikan bahan pertimbangan apakah dokumen yang diuji tersebut merupakan hasil dari meniru karya seseorang atau tidak [1]. Menurut Astuti [3], dalam penelitiannya menyatakan bahwa salah satu kasus yang merebak di negeri ini adalah maraknya berbagai kasus plagiarisme yang terjadi di beberapa perguruan tinggi baik perguruan tinggi negeri maupun perguruan tinggi swasta. Bahkan, Menteri Pendidikan Nasional, Muhammad Nuh menyebut, “maraknya tindak plagiarisme menunjukkan lemahnya pendidikan karakter, budaya, dan moral insan di dunia akademik”, Hal ini menjadi contoh buram dalam potret dinamika pendidikan.

Plagiarisme [4] terjadi dalam karya ilmiah dan tugas akhir mahasiswa dikarenakan memiliki kemiripan yang hampir sama pada karya satu dengan karya lain. Kemiripan sebuah karya ilmiah, karya seni maupun hasil sebuah pencapaian akan mudah bagi seseorang menyebutnya sebagai plagiarisme. Namun, bagi

seorang ahli atau yang menggeluti bidang tersebut akan memiliki beberapa argumen sebelum seseorang menyebutnya sebagai plagiarisme atau memang karya atau hal baru. Kemiripan kata khususnya akan sangat membantu dalam menyimpulkan apakah karya ilmiah atau dokumen ilmiah tersebut masuk dalam kategori plagiarisme atau tidak, karena plagiarisme memiliki syarat-syarat khusus untuk sebuah dokumen atau karya ilmiah untuk bisa masuk dalam kategorinya. Upaya untuk mengatasi masalah kemiripan tidak cukup dengan mengingatkan bahwa tindakan tersebut melanggar hak kekayaan intelektual. Surahman [5], Pencegahan dan pendeteksian merupakan salah satu cara yang paling efektif untuk meningkatkan kualitas pendidikan dan mengurangi kemiripan dokumen satu dengan yang lain. Akan tetapi, pendeteksian sangat sulit jika dilakukan secara manual, sehingga diperlukan suatu sistem cerdas yang mampu mendeteksi kemiripan kata. Pada penelitian Surahman, disarankan untuk mengembangkan Algoritma *Rabin-Karp* berbasis web. Sehingga diperlukan sistem untuk mendeteksi karya lebih mudah dengan menggunakan sistem berbasis web. Agar mempermudah pengguna untuk mengecek karya tersebut melalui internet.

Salah satu metode dalam sistem cerdas yang dapat digunakan yaitu metode string matching dengan menggunakan Algoritma *Rabin-Karp*. Algoritma *Rabin-Karp* adalah multiple pattern search yang sangat efisien untuk mencari string dengan pola banyak dan melakukan proses pendeteksian kemiripan kata dengan menghasilkan nilai persentase kemiripan [6]. Ada beberapa penelitian yang dijadikan referensi yaitu. “Algoritma *Rabin-Karp* berhasil diimplementasikan pada sistem pendeteksi plagiarisme dokumen tugas mahasiswa. Hasil perhitungan akurasi dengan confusion matrix pada sistem pendeteksi plagiarisme dokumen tugas mahasiswa adalah 90% yang diperoleh dari 20 perbandingan dokumen tugas mahasiswa”[7]. Selain itu penggunaan Algoritma *Rabin-Karp* untuk pendeteksian plagiarisme dilakukan oleh Achmad Jauhari [8], pada scenario 1 plagiarisme word to word memiliki tingkat similarity tertinggi pada gram 2 dengan nilai *Similarity* 98,24%, dan pada scenario 2 plagiarisme mozaik nilai *Similarity* tertinggi juga pada gram 2 dengan nilai *Similarity* 98,98 %. Selanjutnya Brinardi Leonardo dan Seng

Hansun [9] Algoritma *Rabin-Karp* lebih efektif dibandingkan dengan Algoritma Jaro-Winkler Distance, karena pada beberapa percobaan Algoritma ini lebih cenderung mendapatkan persentase kemiripan yang lebih tinggi pada dokumen yang diuji dibandingkan dengan Jaro-Winkler Distance.

Seperti yang ditunjukkan oleh hasil studi sebelumnya yang dijelaskan di atas, Algoritma *Rabin-Karp* dapat dianggap sebagai metode yang efektif dan akurat yang digunakan dalam proses pendeteksi kemiripan dokumen. Oleh karena itu, metode *Rabin-Karp* cocok untuk diterapkan dalam studi kasus ini.

1.2 Tujuan dan Manfaat

Tujuan dari penelitian tugas akhir ini adalah:

1. Mengimplementasikan Algoritma *Rabin-Karp* pada sistem pendeteksi kemiripan kata pada dua dokumen dan menghasilkan tingkat persentase kemiripan kata.
2. Sistem memproses dokumen menggunakan teks dan upload dokumen.
3. Menampilkan rata-rata waktu proses eksekusi sistem.

Manfaat dari penelitian tugas akhir ini adalah:

1. Mempermudah dan mempercepat pengecekan dokumen, yang mungkin menyerupai dokumen penelitian yang ada, lebih akurat dibandingkan dengan pengecekan manual.

1.3 Rumusan Masalah

Berdasarkan penjelasan latar belakang di atas, maka dapat dirumuskan masalah yang dihadapi : bagaimana mengimplementasikan Algoritma *Rabin-Karp* untuk sistem pendeteksi kemiripan kata pada dua dokumen.

1.4 Batasan Masalah

1. Sistem dapat mendeteksi kemiripan kata antara dua dokumen skripsi, terdiri dari pendahuluan, abstrak, dan kesimpulan.
2. Dokumen input adalah dokumen yang berformat (*.txt) dan (*.docx).
3. Sistem hanya memproses dokumen teks berbahasa Indonesia.

4. Nilai k pada *parsing K-Gram* ditentukan sebanyak 5 ukuran perkata.
5. Aplikasi ini hanya tersedia pada basis web.

1.5 Metode Penelitian

Dalam hal ini ada beberapa metodologi yang akan digunakan dalam penelitian tugas akhir dengan urutan pembuatan sebagai berikut :

1. Studi Literatur
Mengumpulkan berbagai macam informasi dari buku, jurnal, artikel ataupun internet yang berhubungan dengan judul penelitian.
2. Perancangan Sistem
Menerapkan semua teori-teori penunjang dan komponen yang telah dikumpulkan dalam membangun sebuah aplikasi pendeteksi kata menggunakan Algoritma *Rabin-Karp*.
3. Pengujian dan Analisa
Melakukan pengujian sistem setelah aplikasi selesai dibuat yang akan menghasilkan beberapa data. Setelah itu proses analisa akan dilakukan untuk mengetahui keberhasilan maupun kesalahan implementasi aplikasi yang telah dibuat.
4. Kesimpulan
Membuat kesimpulan yang didapat dari hasil pengujian dan analisa yang telah dilakukan sebelumnya.

1.6 Sistematika Penulisan

Sistematika penulisan yang digunakan dalam tugas akhir ini terbagi menjadi beberapa pokok bahasan, yaitu :

BAB I PENDAHULUAN

Bab ini menjelaskan secara singkat tentang latar belakang masalah, maksud dan tujuan, batasan masalah, rumusan masalah, metode Penelitian, sistematika penulisan dengan maksud memberikan gambaran tentang isi tugas akhir ini.

BAB II TEORI PENUNJANG

Pada bab ini akan di jelaskan mengenai dasar-dasar teori, rujukan dan metode yang berhubungan dengan judul

BAB III ANALISIS DAN PERANCANGAN SISTEM

Bab ini membahas proses pengembangan sistem pada tahap gambaran umum sistem, spesifikasi dan analisis kebutuhan sistem, dengan hasilnya berupa desain dan rancangan sistem yang dikembangkan.

BAB IV HASIL PENGUJIAN DAN ANALISA

Bab ini mencakup hasil pengujian dan analisa dari perancangan sistem yang telah dibuat. Maka akan diketahui sistem tersebut sudah sesuai dengan kebutuhan dengan tujuan yang telah ditentukan.

BAB V KESIMPULAN DAN SARAN

Bab ini berisi tentang kesimpulan dari pembuatan aplikasi Algoritma *Rabin* untuk sistem deteksi kemiripan kata pada dua dokumen serta berisi saran-saran yang ditujukan kepada semua pihak yang bersangkutan.