

BAB I

PENDAHULUAN

1.1. Latar Belakang

Pesatnya perkembangan teknologi membawa banyak manfaat yang amat besar, kemudahan memperoleh informasi salah satunya. Walaupun dengan banyak manfaat yang diberikan oleh perkembangan teknologi tersebut, tetap saja memiliki dampak negative. Kemudahan yang didapat sering disalahgunakan oleh beberapa oknum untuk melakukan kecurangan maupun kejahatan. Plagiarisme merupakan salah satu bentuknya. Inilah yang kemudian menjadi masalah karena merugikan pihak-pihak yang bersangkutan. Pertama, jelas pembuat karena sesuatu yang telah dia teliti kemudian diakui oleh orang lain. Kedua, orang yang mengambil hak orang lain juga merugi dikarenakan dapat mematikan plagiasi bisa mematikan kreatifitas seseorang.

Untuk mendeteksi sebuah karya secara manual jelas dibutuhkan waktu, terlebih jika yang akan diperiksa mencapai ribuan jumlahnya. Beberapa metode untuk menganalisa apakah dokumen termasuk plagiat atau bukan di antaranya pencocokan *substring*, kesamaan kata kunci dan dokumen *fingerprinting*.

Karena itu dibangun sebuah sistem untuk mendeteksi kemiripan dokumen teks. Sistem ini nantinya akan memberi keluaran berupa presentase kemiripan antara dokumen yang dicocokkan. Selanjutnya ini akan menjadi bahan pertimbangan apakah suatu dokumen termasuk plagiat atau bukan.

Dalam tugas akhir ini akan dibandingkan dua tipe pencocokan namun menggunakan satu metode. Satu pencocokan menggunakan tahap *preprocessing*, sedangkan pencocokan yang satunya tidak menggunakan tahap *preprocessing*. Penelitian ini dilakukan untuk mengetahui perbedaan waktu komputasi dan tingkat similaritas pada program yang menggunakan *preprocessing* dan tidak menggunakan *preprocessing*.

Dengan adanya tahap *preprocessing* memang data yang akan diolah semakin sedikit, namun proses yang dilewati lebih banyak, seperti : mengurangi dimensi, menghilangkan *noise* dan memperbaiki struktur kalimat. Tahapan yang dilakukan adalah *case folding*, *tokenizing*, *filtering* dan *stemming*. Jika tahap *preprocessing* dihilangkan, maka proses diatas akan hilang namun jumlah data

yang diproses akan semakin banyak karena tidak mengalami proses apapun sebelum dibandingkan. Setelah keluar hasil dari kedua proses tersebut, akan dibandingkan waktu komputasi tingkat similaritas nya.

1.2. Perumusan Masalah

Adapun rumusan masalah berdasarkan latar belakang tersebut adalah :

- a. Bagaimana mendesain dan membuat aplikasi yang dapat digunakan melakukan pendeteksian plagiasi terhadap dokumen teks ?
- b. Bagaimana menentukan sebuah karya ilmiah termasuk kategori plagiat atau bukan ?
- c. Bagaimana pengaruh penggunaan *preprocessing* terhadap presentase similaritas dan waktu komputasi pada algoritma Rabin-Karp ?

1.3. Tujuan dan Manfaat

a. Tujuan

Tujuan dari Tugas Akhir ini adalah untuk menjawab rumusan masalah. Adapun tujuan tersebut sebagai berikut:

1. Merancang aplikasi pendeteksi plagiat pada karya ilmiah.
2. Mengukur tingkat plagiat suatu karya ilmiah untuk menentukan termasuk kategori plagiat atau bukan.
3. Mengetahui pengaruh *preprocessing* terhadap presentase similaritas dan waktu komputasi pada algoritma Rabin-Karp.

b. Manfaat

Manfaat dari penelitian ini adalah :

1. Dapat membantu sebagai bahan pertimbangan dalam menentukan terjadinya plagiarisme dalam suatu karya ilmiah.
2. Dapat menghasilkan suatu sistem yang dapat mendeteksi apakah suatu karya termasuk plagiat atau bukan.
3. Dapat mengetahui presentase kemiripan (similaritas) antara dokumen asli dan dokumen yang akan diuji.

1.4. Batasan Masalah

Ruang lingkup masalah pada penelitian Tugas Akhir ini adalah sebagai berikut:

- a. Aplikasi ini menerapkan algoritma Rabin-Karp
- b. Hanya menguji data berupa teks, tidak menguji data berupa suara atau gambar.
- c. Data yang diuji bertipe doc dan txt.
- d. Data yang diuji hanya bagian abstrak dari karya ilmiah, bukan seluruhnya.
- e. Data yang diuji menggunakan bahasa Indonesia.

1.5. Metodologi Pengerjaan Tugas Akhir

Secara umum metodologi penelitian dijelaskan sebagai berikut :

1. Studi Pustaka
Studi kepustakaan dengan memahami referensi baik berupa buku, jurnal, artikel dalam bentuk *hardcopy* ataupun *online* yang berhubungan dengan pencocokan kalimat dengan Metode Rabin-Karp.
2. Desain dan Perancangan
Desain dan perancangan sistem deteksi plagiat dilakukan setelah memahami algoritma Rabin-Karp dari studi pustaka yang telah dilakukan.
3. Pembuatan Program
Pembuatan program meliputi pembuatan program pencocokan kalimat dengan algoritma Rabin-Karp.
4. Pengujian deteksi plagiat dan Analisa Program
Pengujian tingkat similaritas dokumen dan waktu komputasi dilakukan dengan membandingkan proses yang menggunakan *preprocessing* dan yang tidak menggunakan *preprocessing*.
5. Penarikan Kesimpulan
Setelah pengujian dilakukan, selanjutnya ditarik kesimpulan tentang performa dari metode yang digunakan.

1.6. Sistematika Penulisan

Sistematika pembahasan dalam Tugas Akhir ini adalah sebagai berikut :

BAB I : PENDAHULUAN

Pada bab ini terdiri dari latar belakang, perumusan masalah, tujuan dan manfaat, batasan masalah, metodologi pengerjaan tugas akhir, dan sistematika penulisan.

BAB II : STUDI PUSTAKA

Pada bab ini dijelaskan mengenai hal yang menjadi pedoman untuk menyelesaikan pengerjaan Tugas Akhir.

BAB III : PERANCANGAN SISTEM

Pada bab ini diuraikan tentang analisis permasalahan, kebutuhan sistem, desain sistem, perancangan, dan pembuatan program.

BAB IV : IMPLEMENTASI SISTEM

Pada bab ini dijelaskan implementasi metode pencocokan kalimat berbasis algoritma Rabin-Karp dan menganalisa tingkat similaritas dan waktu komputasi dari algoritma ini.

BAB V : PENUTUP

Pada bab ini berisi kesimpulan terhadap sistem yang sudah diimplemetasikan, serta saran untuk pengembangan sistem dan penelitian berikutnya.