

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan pada dunia teknologi informasi saat ini telah dirasakan dikalangan mahasiswa. Hal ini menyebabkan penyalahgunaan teknologi informasi yang semulanya digunakan untuk mencari referensi dengan lebih mudah namun kini digunakan sebagai alat untuk duplikasi atau plagiarisme. Faktor yang mempermudah mahasiswa melakukan plagiarisme adalah fasilitas pada komputer yang dapat melakukan menyalin dan mengubah teks antar dokumen. Dalam kenyataannya proses plagiarisme tidak hanya pada tugas essay saja tetapi pada kode program [1].

Sebelum menentukan suatu tugas mahasiswa termasuk melakukan plagiarisme atau tidak termasuk melakukan plagiarisme, maka sangatlah penting untuk mengetahui apa definisi plagiat atau plagiarisme. Plagiat adalah pengambilan karangan, pendapat, dan sebagainya dari orang lain dan menjadikannya seakan – akan karangan / pendapat sendiri, misalnya menerbitkan karya tulis orang lain atas nama dirinya sendiri. Orang yang melakukan plagiat disebut plagiator atau penjiplak. Berdasarkan pernyataan ini maka kesamaan atau kemiripan isi dari tugas antara mahasiswa termasuk tindakan plagiarisme [1]. Saat ini yang sering terjadi di perguruan tinggi untuk melakukan pengecekan adanya tindakan plagiarisme pada dokumen tugas mahasiswa dilakukan secara manual dan membandingkan dengan

dokumen tugas yang lain. Dengan begitu, pengecekan plagiarisme akan memakan waktu yang sangat lama.

Metode yang dapat digunakan adalah metode pencocokan string yaitu salah satunya menggunakan *Algoritma Rabin-Karp*. *Algoritma Rabin-Karp* adalah salah satu algoritma pencarian string dikembangkan oleh *Michael O. Rabin* dan *Richard M. Karp* pada tahun 1987 yang menggunakan fungsi hashing untuk menemukan *pattern* di dalam string teks [2]. *Algoritma Rabin-Karp* merupakan algoritma untuk pencocokan string multi *pattern*. Pada pencocokan string multi *pattern* paket atau informasi yang dicari berdasarkan beberapa susunan pola string. *Algoritma Rabin-Karp* melakukan pergeseran dari kiri ke kanan, fungsi dari algoritma *Rabin-Karp* menghasilkan efisiensi waktu yang baik dalam pencarian string yang memiliki lebih dari satu pola [3].

Berdasarkan uraian diatas, peneliti akan melakukan implementasi algoritma *Rabin-Karp* ke dalam sistem berbasis web agar dapat mendeteksi seberapa besar persentase plagiarisme atau kemiripan antar tugas mahasiswa.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan, maka didapatkan rumusan masalah yaitu “bagaimana algoritma *Rabin-Karp* dapat di implementasikan pada sistem dan mampu mendeteksi tingkat kemiripan tugas antar mahasiswa ?”.

1.3 Batasan Masalah

Adapun batasan masalah pada penelitian ini adalah sebagai berikut :

1. Fokus dari penelitian ini adalah menerapkan algoritma *Rabin-karp* pada sistem berbasis web untuk mendeteksi tingkat kemiripan tugas mahasiswa.
2. Dokumen teks yang dapat diproses oleh sistem adalah dokumen yang menggunakan format .doc, .docx dan .pdf.
3. Jumlah berkas yang mampu ditangani oleh sistem adalah 20 berkas dengan maksimal 30 halaman per berkas.
4. Nilai k pada parsing k -gram ditentukan sebanyak 4 ukuran perkata.

1.4 Tujuan Penelitian

Tujuan penelitian dalam skripsi ini adalah menyajikan sistem yang dapat menampilkan persentase tingkat kemiripan tugas antar mahasiswa menggunakan algoritma *Rabin-Karp*.

1.5 Manfaat Penelitian

Hasil penelitian ini diharapkan dapat bermanfaat untuk :

1. Peneliti
Meningkatkan kemampuan dalam *text mining* dan pemrograman.
2. Perguruan Tinggi
Mengetahui persentase kemiripan tugas antar mahasiswa yang dapat digunakan sebagai bukti melakukan plagiarisme.

1.6 Metode Penelitian

1. Pengumpulan Data

a. Studi Pustaka

Studi pustaka dilakukan untuk mempelajari penelitian-penelitian dengan tema sejenis yang pernah dilakukan sebelumnya. Penelitian-penelitian ini dijadikan referensi pada penelitian ini. Sumber yang dipakai dapat berupa buku, jurnal, *paper*, *website*, dan lain-lain. Pada penelitian ini, referensi yang dicari adalah pendeteksian kemiripan antar dokumen.

b. Pengumpulan Data

Data *input* yang akan dipakai pada penelitian ini berupa dokumen tugas mahasiswa dalam format *.doc, *.docx dan *.pdf.

2. Analisis dan Perancangan

Pada tahap ini akan dilakukan analisis sistem untuk mengetahui kebutuhan sistem yang akan dibuat. Setelah itu, melakukan perancangan sistem untuk mendeteksi kemiripan dokumen tugas antar siswa.

3. Implementasi

Implementasi dilakukan dengan menerapkan rancangan sistem yang telah dibangun sebelumnya. Implementasi sistem ini menggunakan *library* Sastrawi untuk *stemming*, algoritma *Rabin-Karp*, bahasa pemrograman PHP, dan database Mysql.

4. Pengujian Sistem

Pada tahap ini akan dilakukan pengujian sistem deteksi plagiarisme tugas mahasiswa menggunakan metode *black box* dan pengujian akurasi algoritma dengan *confusion matrix*.

5. Penulisan Laporan

Hasil penelitian kemudian dituliskan ke dalam sebuah laporan. Dalam penulisan laporan penelitian akan ditarik pula suatu kesimpulan berdasarkan hasil implementasi dan pengujian, serta saran untuk penelitian berikutnya.

1.7 Sistematika Penulisan

Rincian sistematika penulisan yang digunakan dalam penelitian ini adalah sebagai berikut:

1. BAB I PENDAHULUAN

Bab ini memuat paparan latar belakang penelitian, rumusan masalah, batasan masalah dalam penelitian, tujuan penelitian, manfaat penelitian, metodologi penelitian, dan sistematika penulisan.

2. BAB II LANDASAN TEORI

Bab ini berisi pembahasan tinjauan pustaka tentang penelitian sebelumnya yang dijadikan referensi pada penelitian ini, serta dasar-dasar teori yang digunakan dalam penelitian ini.

3. BAB III ANALISIS DAN PERANCANGAN

Bab ini berisi analisis dari sistem yang akan dibuat, serta gambaran singkat tentang tahap-tahap pembuatan sistem tersebut.

4. BAB IV IMPLEMENTASI DAN PENGUJIAN

Bab ini berisi implementasi sistem yang telah dibuat menggunakan *library*, bahasa pemrograman dan database yang telah ditentukan serta pengujian sistem menggunakan metode *black box* dan pengujian akurasi algoritma dengan *confusion matrix*.

5. BAB V PENUTUP

Bab ini membahas kesimpulan dari penelitian serta saran untuk penelitian selanjutnya.

