

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Sesuai fitrahnya manusia menginginkan kemudahan dalam segala hal. Jika tidak diimbangi dengan mental positif untuk berkreasi menciptakan karya *original*, sikap ini akan membawa kepada tindakan yang merugikan diri sendiri dan orang lain, misalnya plagiarisme.

Plagiarisme selalu menjadi perhatian di banyak sektor, terutama sektor akademis. Sering kita temui pekerjaan teman yang dihargai dengan nilai "0" karena diketahui telah menyalin hasil pekerjaan temannya yang lain. Praktik menyalin beberapa bagian atau keseluruhan tulisan tanpa mencantumkan sumber menjadi hal biasa yang sering ditemukan pada tugas harian, makalah, bahkan skripsi.

Pencegahan dan pendeteksian merupakan cara yang dapat dilakukan untuk mengurangi plagiarisme. Pencegahan berarti menghalangi munculnya plagiarisme yang lebih ditekankan kepada moral masyarakat dan sistem pendidikan. Cara ini akan memberikan efek jangka panjang yang luar biasa. Pendeteksian berarti cara untuk mengungkapkan plagiarisme.

Pendeteksian dapat dilakukan dengan teknik manual, hanya saja akan memerlukan waktu yang sangat lama untuk membandingkan satu dokumen dengan dokumen lainnya yang ada di *database*. Apalagi jika plagiat hanya mengutip sebagian dari artikel atau menggabungkan potongan artikel, akan sangat susah untuk menafsirkan gaya penulisannya.

Sejauh ini sudah ada beberapa perangkat lunak yang didesain untuk mendeteksi plagiarisme dokumen, diantaranya yang paling terkenal adalah Turnitin, Eve2, CopyCatchGold, WordCheck, Glatt, Moss, JPlag. Berdasarkan analisis informasi yang ada di web, pendeteksi terbaik sesuai fungsinya adalah Turnitin (Lukashenko, dkk, 2007).

Pendeteksian plagiarisme atau plagiat menggunakan konsep *similarity* atau kemiripan dokumen merupakan salah satu cara untuk mendeteksi *copy&paste plagiarism* dan *disguised plagiarism*. Menggunakan metode *fingerprinting*, pendeteksian plagiat ini tidak memperhatikan adanya penulisan sumber rujukan.

Pada sistem deteksi ini akan diaplikasikan *text mining* dan algoritma Rabin Karp sekaligus. *Text mining* sendiri akan diaplikasikan untuk tahap *preprocessing*. Algoritma Rabin Karp adalah algoritma *multiple pattern search* yang sangat efisien untuk mencari *string* dengan pola banyak. Konsep kerjanya adalah mencari *string* yang sesuai dengan *string* masukan (dokumen yang dibandingkan) setelah melalui proses *hashing*.

Skripsi ini membahas secara skematis bagaimana algoritma Rabin Karp bekerja sekaligus implementasinya pada sebuah aplikasi dalam mendeteksi plagiarisme dokumen.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang diatas, dirumuskan permasalahan sebagai berikut :

1. Bagaimana membangun sebuah sistem yang dapat mendeteksi plagiat pada dokumen teks?

2. Bagaimana implementasi algoritma Rabin Karp dalam pendeteksian plagiat?
3. Bagaimana performa dan akurasi algoritma Rabin Karp dalam mendeteksi plagiat?

### 1.3 Batasan Masalah

Batasan masalah dalam skripsi ini diantaranya :

1. Sistem hanya menguji data teks, tidak menguji data gambar maupun tabel.
2. Sistem hanya mendeteksi kemiripan dokumen dan tidak memperhatikan adanya penulisan sumber/ rujukan.
3. Sistem hanya mendeteksi dokumen-dokumen internal dalam *database*, tidak membandingkan dengan dokumen eksternal di internet.
4. Sistem hanya menguji data teks berekstensi *.txt*.
5. Hanya menguji teks berbahasa Indonesia.
6. Sistem tidak memperhatikan padanan kata/ sinonim.
7. Sistem tidak memperhatikan kesalahan ejaan dan penulisan.

### 1.4 Tujuan Penelitian

Adapun tujuan dari pembuatan skripsi ini yaitu:

1. Merancang aplikasi untuk mendeteksi plagiat pada dokumen teks dengan konsep *similarity* menggunakan algoritma Rabin Karp.
2. Membuktikan akurasi algoritma Rabin Karp dalam mendeteksi *local similarities* pada identifikasi dokumen.
3. Mengetahui prosentase kemiripan (*similarity*) dokumen yang diuji dibandingkan dokumen asli.

4. Meneliti performa algoritma Rabin Karp beserta faktor-faktor yang mempengaruhi sistem dalam meningkatkan keakuratan deteksi.

### 1.5 Manfaat Penelitian

Manfaat yang diharapkan dari penyusunan skripsi ini antara lain:

1. Membantu institusi dan para pengajar dalam mengecek hasil pekerjaan peserta didik.
2. Sebagai alat bantu pertimbangan dalam menentukan kemiripan antardokumen untuk kemudian merumuskan solusi atas tindakan tersebut.
3. Sebagai alat motivasi agar setiap orang dapat lebih *fair* dalam menciptakan sebuah karya.

### 1.6 Metodologi Penelitian

Metodologi yang digunakan dalam penyusunan skripsi ini yaitu:

1. Studi Literatur

Pengumpulan data melalui berbagai literatur, diantaranya dengan mempelajari artikel-artikel dan jurnal-jurnal terkait plagiat, *document similarity*, algoritma Rabin Karp, *text mining*, *information retrieval* serta mempelajari artikel-artikel terkait dari internet.

2. Observasi

Melakukan pengamatan dan pengujian terhadap beberapa aplikasi pendeteksi plagiat di internet sehingga mempunyai gambaran tentang sistem yang telah terimplementasikan.

### 3. Analisis dan Perancangan Sistem

Melakukan analisis kebutuhan dan perancangan sistem serta menguji algoritma terhadap data dengan melakukan perhitungan manual.

### 4. Implementasi Sistem

Mengimplementasikan perancangan sistem ke dalam bentuk aplikasi pendeteksian plagiat.

### 5. Uji coba dan Evaluasi

Melakukan uji coba sistem kemudian melakukan evaluasi untuk memperbaiki kekurangannya.

## 1.7 Sistematika Penulisan

Laporan penelitian ini akan disusun secara sistematis kedalam 5 bab, masing-masing bab akan diurutkan sebagai berikut :

### BAB I PENDAHULUAN

Pada bab ini akan diuraikan mengenai latar belakang masalah, perumusan masalah, batasan masalah, tujuan, manfaat, metodologi, dan sistematika penulisan penelitian.

### BAB II LANDASAN TEORI

Pada bab ini akan diuraikan tentang teori yang digunakan sebagai dasar dalam analisis penelitian. Mencakup *text processing*, *document similarity*, *text mining*, *string matching*, dan teknik *hashing*.

### BAB III ANALISIS DAN PERANCANGAN SISTEM

Pada bab ini akan diuraikan tentang metode input data, perancangan program, dan proses analisis.

#### BAB IV IMPLEMENTASI DAN PEMBAHASAN

Pada bab ini membahas implementasi metode dan pengujian sistem deteksi plagiarisme.

#### BAB V PENUTUP

Bab ini berisi kesimpulan yang didapat dari hasil analisis, perancangan, dan implementasi sistem, serta saran untuk pengembangan lebih lanjut

