



Hak cipta dan penggunaan kembali:

Lisensi ini mengizinkan setiap orang untuk menggubah, memperbaiki, dan membuat ciptaan turunan bukan untuk kepentingan komersial, selama anda mencantumkan nama penulis dan melisensikan ciptaan turunan dengan syarat yang serupa dengan ciptaan asli.

Copyright and reuse:

This license lets you remix, tweak, and build upon work non-commercially, as long as you credit the origin creator and license it on your new creations under the identical terms.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Dewasa ini, perkembangan teknologi yang pesat dan penyebaran internet yang cepat sangat berpengaruh terhadap gaya hidup manusia. Pertukaran informasi antara satu pihak dengan pihak lain kini lebih mudah dilakukan dengan menggunakan internet. Tentu ini berdampak positif bagi masyarakat. Namun, fakta tersebut juga dapat memberikan dampak negatif. Internet memfasilitasi penggunaanya agar bisa mengakses berbagai macam konten dan informasi dari seluruh dunia. Kemudahan dalam mengakses konten dan informasi, khususnya berupa tulisan, dapat menjadi salah satu faktor bagi manusia untuk melakukan tindakan penggandaan dokumen secara langsung, tanpa memasukkan narasumber. Salah satu contoh yang umum adalah dalam membuat tugas akhir, seluruh isinya merupakan hasil duplikat dari tugas akhir milik orang lain. Dengan demikian, terjadilah suatu tindakan yang disebut dengan plagiarisme.

Kata plagiarisme sudah tidak asing lagi bagi banyak orang. Bukan hanya karena sering disebutkan secara lisan maupun tulisan, tetapi kasus yang berhubungan dengan plagiarisme sering terjadi. Plagiarisme merupakan suatu tindakan dimana seseorang mengakui bahwa karyanya adalah buatan sendiri tetapi pada kenyataannya itu adalah karya milik orang lain (Kirszner & Mandell, 2008). Adanya praktik plagiarisme membawa dampak buruk tidak hanya bagi karya yang dijiplak, tetapi juga bagi yang menjiplak. Segala sesuatu menjadi serba instan karena yang dilakukan hanyalah penggandaan. Ini mendorong manusia untuk

bermalas-malasan dan tidak mengasah kreatifitas (Parsons, 2007). Oleh karena itu, praktik plagiarisme harus secara perlahan dicegah sejak dini.

Untuk mengatasi praktik plagiarisme, khususnya dalam dunia akademik, mengingatkan secara lisan dan tulisan kepada mahasiswa untuk tidak melakukan tindakan plagiarisme saja tidaklah cukup. Pendeteksian praktik plagiarisme terhadap dokumen teks merupakan sebuah solusi yang baik. Akan tetapi, proses pendeteksian secara manual akan sangat sulit dan memakan waktu untuk dilakukan karena jumlah teks yang banyak. Inilah mengapa dibutuhkan suatu aplikasi yang bisa mendeteksi plagiarisme (Poon, Sugiyama, Tan, & Kan, 2012).

Aplikasi pendeteksi plagiarisme pada dokumen teks bisa dikembangkan dengan menggunakan algoritma pencocokan teks atau yang biasa disebut dengan *string matching*. Metode ini dapat digunakan untuk menghitung kemiripan teks antara satu dokumen dengan dokumen lainnya (Osman, Salim, & Abuobieda, 2012). Banyak algoritma pencocokan *string* yang ada, salah satunya adalah algoritma Rabin-Karp yang akan dipakai dalam aplikasi pendeteksi plagiarisme ini.

Algoritma Rabin-Karp didasari dengan teknik *hashing*. Setiap pola *string* dikonversi menjadi sebuah nilai *hash*. Dari nilai-nilai tersebut, akan dilakukan pencocokan *string* yang dicari dengan teks (Puntambekar, 2009). Kelebihan dari algoritma Rabin-Karp dibandingkan algoritma pencocokan *string* lainnya adalah kemampuan dalam mencari banyak pola *string* (Cormen, Leiserson, Rivest, & Stein, 2003). Di samping itu, terdapat algoritma Rabin-Karp yang sudah

dimodifikasi. Dengan begitu, bisa dibandingkan tingkat kemiripan dan waktu proses antara dua algoritma (Chillar & Kochar, 2008).

1.2 Rumusan Masalah

Masalah yang dapat dirumuskan dalam penelitian ini adalah

1. Bagaimana cara merancang suatu aplikasi yang dapat mendeteksi plagiarisme pada dokumen teks ?
2. Bagaimana algoritma Rabin-Karp bisa digunakan untuk mendeteksi plagiarisme pada dokumen teks ?
3. Bagaimana pengaruh modifikasi algoritma Rabin-Karp terhadap persentase kemiripan dan waktu proses ?

1.3 Batasan Masalah

Beberapa batasan yang ditentukan dalam penelitian ini yaitu sebagai berikut.

1. Bahasa yang digunakan adalah bahasa Indonesia yang baku berdasarkan EYD.
2. Data yang diuji berupa teks, tidak menguji data berupa suara atau pun gambar.
3. Data yang akan diuji berformat .txt, .pdf, .doc, dan .docx.
4. Aplikasi tidak memperhatikan kesalahan dalam pengejaan, sinonim, dan *grammar*.
5. Nilai *n-grams* adalah antara 1 dan 5.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah

1. Membuat suatu aplikasi yang mana dapat mendeteksi plagiarisme pada dokumen teks.
2. Mengimplementasikan algoritma Rabin-Karp dalam aplikasi pendeteksi plagiarisme pada dokumen teks.
3. Membuktikan bahwa modifikasi yang dilakukan pada algoritma Rabin-Karp memiliki pengaruh terhadap waktu proses dan persentase kemiripan.

1.5 Manfaat Penelitian

Diharapkan penelitian ini dapat memberikan manfaat sebagai berikut.

1. Hasil aplikasi dapat digunakan sebagai bahan pertimbangan dalam menentukan plagiarisme dalam dokumen teks berbahasa Indonesia.
2. Aplikasi dapat digunakan untuk menghitung persentase kemiripan antara dokumen asli dengan dokumen uji.

1.6 Sistematika Penulisan

Sistematika penulisan laporan skripsi ini dijelaskan sebagai berikut.

Bab I Pendahuluan

Berisi latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan skripsi.

Bab II Tinjauan Pustaka

Berisi landasan teori mengenai plagiarisme, dokumen elektronik, Bahasa Indonesia, pencocokan *string*, *text mining*, algoritma Rabin-Karp, *hashing*, *n-grams*, dan nilai *similarity*.

Bab III Analisis dan Perancangan Aplikasi

Berisi analisis permasalahan serta solusinya, spesifikasi umum, dan rancangan aplikasi yang akan dibuat untuk mendeteksi plagiarisme pada dokumen teks berbahasa Indonesia.

Bab IV Implementasi dan Uji Coba

Berisi penjelasan mengenai implementasi dan hasil uji coba aplikasi yang didasarkan dengan rancangan pada bab sebelumnya.

Bab V Simpulan dan Saran

Berisi simpulan dan saran dari penelitian yang dapat digunakan untuk penelitian selanjutnya.

UMMN