

# 1. Pendahuluan

## 1.1 Latar belakang

Semakin berkembangnya teknologi komputerisasi, semakin mudah seseorang menyelesaikan pekerjaannya atau menghasilkan sesuatu. Dengan menggunakan komputer, seseorang dapat menciptakan sebuah hasil karya berdasarkan referensi-referensi dari internet atau yang telah tersimpan dalam media penyimpanan komputer tersebut. Namun kemudahan ini memicu tindakan yang seharusnya dihindari menjadi semakin mudah dilakukan, yaitu tindakan plagiarisme.

Plagiarisme adalah penjiplakan atau pengambilan karangan, pendapat, dan sebagainya, dari orang lain dan menjadikannya seolah karangan dan pendapat sendiri[1]. Plagiarisme banyak terjadi dalam pembuatan teks, baik berupa tugas-tugas akademis atau artikel-artikel yang dipublikasikan dalam internet. Plagiarisme seperti ini semakin mudah dilakukan dengan adanya fasilitas *copy-and-paste* dalam komputer.

Untuk mencegah plagiarisme teks, banyak upaya yang dilakukan, salah satunya dengan memeriksa hasil karya seseorang secara manual, sebelum dinilai atau dipublikasikan. Jika sebuah teks dideteksi sebagai hasil plagiat, bukan hanya teks tersebut tidak akan dinilai dan dipublikasikan, tetapi pembuat teks juga dihukum sesuai ketentuan-ketentuan yang berlaku. Hal ini dilakukan agar timbul efek jera pada diri pelaku dan untuk mencegah orang lain melakukan tindakan yang sama.

Akan tetapi pemeriksaan teks secara manual untuk mendeteksi plagiarisme bisa menjadi pekerjaan yang sulit jika terdapat banyak teks yang harus diperiksa. Dibutuhkan waktu yang cukup lama untuk melakukan pemeriksaan seluruh teks tersebut. Oleh karena itu beberapa lembaga mencoba mengembangkan sebuah sistem yang dapat mendeteksi plagiarisme secara otomatis.

Namun sistem yang dapat mendeteksi plagiarisme seperti ini cukup sulit untuk direalisasikan. Hal ini karena adanya masalah-masalah hukum yang berkaitan dengan *fair use*, yang membolehkan penggunaan bahan-bahan yang telah dilindungi hak cipta dalam karya penulis lain dengan syarat-syarat tertentu [2].

Berdasarkan alasan tersebut, sistem seperti ini lebih tepat dianggap sebagai alat bantu pendeteksian plagiarisme. Disebut alat bantu karena sistem ini hanya mendeteksi kemiripan-kemiripan antara teks yang dibandingkan. Sedangkan pengkategorian plagiarisme dilakukan secara manual berdasarkan kemiripan-kemiripan tersebut.

Tugas akhir ini menganalisa pembuatan sistem yang berupa alat bantu pendeteksian plagiarisme teks dengan menerapkan algoritma Rabin-Karp. Algoritma ini adalah algoritma yang dapat menemukan posisi beberapa *string* berukuran sama dalam *string* yang lebih besar. Karena teks juga dapat dianggap sebagai *string*, maka algoritma ini dapat diterapkan sebagai bagian dari alat bantu pendeteksian plagiarisme teks. Dengan penerapan algoritma ini, diharapkan sistem dapat menemukan kemiripan-kemiripan dalam teks, sehingga mempermudah pendeteksian plagiarisme secara manual.

## 1.2 Perumusan masalah

Permasalahan yang muncul dan akan dianalisa dalam tugas akhir ini adalah sebagai berikut:

1. Masukan algoritma Rabin-Karp adalah beberapa *string* (himpunan *string*) berukuran sama dan sebuah *string* lain. Bagaimana membuat sebuah teks menjadi sebuah himpunan *string* berukuran sama sehingga dapat dijadikan sebagai masukan algoritma Rabin-Karp.
2. Tidak ada ketentuan khusus untuk ukuran *string* dalam himpunan *string*. Bagaimana menentukan ukuran *string* yang baik, karena berdasarkan analisa awal, algoritma Rabin-Karp mungkin menghasilkan keluaran yang berbeda untuk ukuran *string* yang berbeda.
3. Keluaran algoritma Rabin-Karp adalah posisi-posisi setiap himpunan *string* dalam *string* lain. Bagaimana mengolah keluaran ini untuk dapat menemukan kemiripan antara teks yang dibandingkan.
4. Berdasarkan hasil analisa awal, terdapat beberapa perubahan dalam teks yang umumnya dilakukan pelaku plagiarisme (dijelaskan dalam bab 2). Bagaimana sistem menemukan perubahan tersebut dalam teks sehingga dapat menentukan apakah sebuah teks memiliki kemiripan dengan teks lain.

Adapun yang menjadi batasan masalah dalam tugas akhir ini adalah sebagai berikut:

1. Sistem hanya difokuskan untuk menemukan kemiripan dalam teks-teks yang ditulis dalam bahasa Indonesia.
2. Masukan sistem adalah dua buah teks yang berupa *file* berekstensi '*.txt*'.
3. Sistem tidak memeriksa kesalahan penulisan kata dalam teks.
4. Sistem tidak memeriksa ada tidaknya *stopword* dalam teks. *Stopword* tetap diikutsertakan dalam penentuan kemiripan.
5. Selain algoritma Rabin-Karp dan metode-metode yang mendukung penggunaannya (diantaranya *hashing*, *sorting*, dan *searching*), serta metode yang dibuat untuk mengolah keluaran algoritma menjadi keluaran sistem, dalam sistem ini tidak diterapkan metode-metode spesifik lain yang mungkin dapat digunakan dalam menemukan kemiripan diantara teks.

## 1.3 Tujuan

Tujuan yang ingin dicapai dalam tugas akhir ini adalah:

1. Menganalisa pengaruh penggunaan ukuran *string* yang berbeda terhadap keluaran sistem, dan menentukan ukuran *string* yang menghasilkan keluaran sistem yang optimal.
2. Menentukan perubahan-perubahan apa saja (poin 4 perumusan masalah) yang dapat ditemukan sistem sehingga dapat menentukan apakah dua teks yang dibandingkan memiliki kemiripan.
3. Menentukan batas tingkat kemiripan yang menjadi acuan sistem dalam menentukan kemiripan (dijelaskan dalam bab 3), sehingga dengan batas tingkat kemiripan ini sistem dapat digunakan dalam berbagai kondisi kemiripan dalam teks.
4. Menganalisa apakah kemiripan yang ditemukan sistem serupa dengan hasil pemeriksaan manual.

## **1.4 Metodologi penyelesaian masalah**

Metodologi penyelesaian masalah pada tugas akhir ini adalah sebagai berikut:

1. Studi Literatur  
Menghimpun dan mempelajari teori-teori mengenai plagiarisme, seperti apa pada umumnya teks hasil plagiarisme, bagaimana mendeteksinya, dan bagaimana cara kerja algoritma Rabin-Karp.
2. Perancangan sistem  
Merancang cara kerja sistem memeriksa teks dan menemukan kemiripan dengan teks lain, dan bagaimana sistem menyajikan hasil pemeriksaan tersebut.
3. Implementasi sistem  
Membangun sistem sesuai hasil perancangan agar dapat melakukan pengujian dan menganalisa hasil pengujian tersebut.
4. Pengujian dan penganalisaan hasil pengujian  
Pengujian dilakukan dengan memeriksa apakah dua teks memiliki kemiripan dengan menggunakan sistem dan secara manual. Hasil pengujian tersebut kemudian dianalisa agar dapat ditarik kesimpulan dari hasil pengujian sistem .
5. Penyimpulan hasil penelitian  
Kesimpulan diambil dari analisis hasil pengujian. Dengan kesimpulan ini, tujuan-tujuan dari penganalisaan dan pengimplementasian algoritma Rabin-Karp dalam alat bantu pendeteksian plagiarisme teks dapat tercapai.
6. Penyusunan laporan  
Laporan digunakan untuk memaparkan analisis, hasil-hasil pengujian, serta kesimpulan yang telah diambil dari analisis dan hasil-hasil pengujian tersebut.