

BAB I

PENDAHULUAN

1.1. Latar Belakang

Di zaman perkembangan teknologi seperti saat ini, informasi dan pengetahuan secara pesat berkembang dimana-mana, kita disugahi banyaknya informasi yang beredar, dengan pesatnya informasi yang beredar akan memudahkan kita dalam mendapatkan informasi. Kemudahan informasi yang didapatkan, dapat menyebabkan kesempatan untuk terjadinya tindakan penjiplakan atau plagiarisme suatu karya.

Menurut (Aziz et al., 2015) “Plagiarisme atau sering disebut plagiat adalah penjiplakan atau pengambilan karangan, pendapat, dan sebagainya dari orang lain dan menjadikannya seolah karangan dan pendapat sendiri”. Plagiat atau biasa disebut dengan penjiplakan adalah sebuah masalah yang sering terjadi di dunia pendidikan, contohnya dalam pembuatan laporan, jurnal, tugas akhir dan karya ilmiah ini sangat rawan terjadinya tindakan *copy-paste* referensi dari internet tanpa mencantumkan sumber aslinya.

Sebagai langkah awal pencegahan tindakan plagiarisme dalam dokumen karya ilmiah diperlukan cara mendeteksi kemungkinan terjadinya tindakan plagiarisme dengan mencocokkan serta membandingkan karya ilmiah tersebut dengan karya ilmiah sebelumnya yang telah dipublikasikan. Untuk mengetahui adanya kemiripan atau kecocokan dokumen karya ilmiah dapat menggunakan algoritma *string matching*. *String matching* yang dalam bahasa Indonesia dikenal dengan istilah pencocokan *string* merupakan suatu algoritma yang digunakan untuk melakukan pencarian semua kemunculan *string* pendek pattern [0.....n-1] yang disebut *pattern* di *string* yang lebih panjang teks [0.....m-1] yang disebut teks (Wulan, 2011).

Beberapa algoritma dapat diterapkan untuk proses *string matching* ini, salah satunya yaitu algoritma Knuth Morris Pratt. Algoritma Knuth Morris Pratt merupakan jenis *exact string matching algorithm*, yang melakukan pencocokan *string* secara tepat dengan susunan karakter dalam *string* yang dicocokkan

memiliki jumlah maupun urutan karakter dalam *string* yang sama (Indriyono, 2018).

Berdasarkan penjelasan di atas, penulis akan melakukan implementasi algoritma Knuth Morris Pratt pada sistem pendeteksi plagiasi dokumen. Dalam penelitian yang dilakukan ini menghasilkan sebuah perangkat lunak aplikasi yang dapat digunakan untuk melakukan pencocokan *string* pada dokumen. Prosedur yang dijalankan dalam penelitian ini adalah program akan melakukan pencarian *string* kemudian membandingkan antara *string* pada dokumen dataset yang terdapat dalam database dengan *string* pada dokumen yang diuji. Apabila pada saat pencarian dan pencocokan ditemukan *string* yang sama, maka selanjutnya *string* pada dokumen uji akan diberi *highlight* warna merah.

Dengan dilakukannya penelitian ini diharapkan nantinya dapat mengetahui hasil implementasi algoritma Knuth Morris Pratt pada sistem pendeteksi plagiasi dokumen, sehingga dapat mengetahui apakah dokumen tersebut terindikasi adanya tindakan plagiarisme atau tidak.

1.2. Rumusan Masalah

Berdasarkan latar belakang permasalahan di atas, dapat diambil rumusan masalah pada penelitian yang dilakukan adalah bagaimana implementasi algoritma Knuth Morris Pratt pada sistem pendeteksi plagiasi dokumen.

1.3. Tujuan

- a. Untuk memenuhi syarat tugas akhir studi S-1 Program Studi Informatika Universitas Muhammadiyah Ponorogo.
- b. Untuk melakukan pendeteksian plagiasi pada dokumen.
- c. Untuk mengetahui hasil implementasi algoritma Knuth Morris Pratt pada sistem pendeteksi plagiasi dokumen.

1.4. Manfaat

Adapun manfaat dari penelitian ini yaitu mengetahui hasil pencocokan string antara dokumen uji dengan dokumen dataset menggunakan implementasi algoritma Knuth Morris Pratt pada sistem pendeteksi plagiasi dokumen.

1.5. Batasan Masalah

Batasan Masalah Pada penelitian ini yaitu:

- a. Menggunakan algoritma Knuth Morris Pratt yang merupakan *exact string matching*.
- b. Pengujian pencocokan *string* pada dokumen dilakukan dengan mendeteksi dokumen dengan dokumen dataset pada database sistem yang akan dibangun.
- c. Sistem pendeteksi plagiasi pada algoritma Knuth Morris Pratt akan dibangun menggunakan Bahasa pemrograman Python.
- d. Sumber dokumen yang akan dibandingkan dan dijadikan acuan dalam sistem merupakan dokumen berupa karya ilmiah yang ditambahkan dalam database sistem.
- e. Dokumen naskah yang akan diuji dalam format berekstensi pdf dan doc.

