

BAB 1

PENDAHULUAN

1.1. Latar Belakang

Di dalam perkuliahan, terutama jenjang akademis S1, mahasiswa dituntut untuk menyelesaikannya dengan cara membuat penelitian dan menyajikannya dalam bentuk skripsi. Dengan kemajuan teknologi, mahasiswa dapat mengakses materi - materi skripsi dengan mudah, namun kadangkala mahasiswa menyalin materi tersebut secara keseluruhan atau sebagian tanpa memberikan keterangan sumber, baik disengaja ataupun tidak, dengan kata lain menjiplak pekerjaan orang lain yang disebut dengan plagiarisme. Akan tetapi, kegiatan plagiarisme adalah melanggar ketentuan penilaian skripsi dan juga hukum, dengan hukuman diantaranya adalah pencabutan gelar, atau kurungan penjara 2 tahun dan denda paling banyak dua ratus juta rupiah, atau untuk mahasiswa adalah teguran, peringatan tertulis, pembatalan nilai hingga pemberhentian dengan hormat dari status sebagai mahasiswa atau pembatalan ijazah apabila telah lulus. Dampak yang ditimbulkan oleh plagiarisme ini adalah kehilangan kreativitas dan orisinalitas dalam membuat suatu karya dan juga bagi pembuat karya adalah kehilangan uang yang dihasilkan dari karya-karya mereka[1].

Plagiarisme adalah kegiatan menyalin ide ataupun tulisan orang lain secara sengaja maupun tidak sengaja tanpa menuliskan sumber dan mengakui bahwa ide atau tulisan tersebut adalah murni dari dirinya[2]. Algoritma pencarian string merupakan salah satu keas algoritma dalam lingkup string

yang mencari letak suatu string pada kumpulan string yang lebih besar Adapun pada teks[3].

Algoritma adalah kumpulan metode atau perintah yang digunakan untuk memecahkan masalah secara terstruktur dan sistematis [4]. Algoritma Rabin-Karp adalah salah satu algoritma pencarian string dikembangkan oleh Michael O. Rabin dan Richard M. Karp pada tahun 1987 dengan memiliki beberapa karakteristik yaitu menggunakan K-Gram dan fungsi hash[5]. Cara kerja algoritma ini adalah string akan diubah sedemikian hingga menjadi sebuah larik string yang Panjang disebut dengan token, kemudian token tersebut diubah ke dalam nilai integer yang disebut nilai hash. Algoritma ini memiliki prinsip jika dua string adalah sama, maka nilai hash dari dua string harus sama, sehingga algoritma hanya akan mencari string yang memiliki nilai hash yang sama dengan nilai string pada pola[6]. Algoritma rabin-karp banyak digunakan dalam studi kasus pencarian mencari sebuah string ataupun pendeteksi plagiarisme pada dokumen.

Alat uji coba algoritma ini adalah dengan membuat sebuah laman web yang didalamnya dimasukkan Algoritma Rabin-Karp, kemudian dimasukkan input-input yang berupa string yang akan di analisa oleh web dan dimunculkan menjadi presentase berapa persen kesamaan antar dua string tersebut. Website adalah sebuah kumpulan laman yang berisi informasi berupa teks, video, suara, maupun gambar yang dapat diakses oleh banyak orang melalui media internet baik dari komputer maupun handphone. Website dapat dibuka dengan menuliskan alamat website ataupun mencari pada laman situs pencarian[7]. Berdasarkan hal-hal tersebut, maka peneliti tertarik untuk melakukan penelitian

dengan judul : “STRING MATCHING MENGGUNAKAN ALGORITMA RABIN KARP UNTUK PENCOCOKAN JUDUL SKRIPSI MAHASISWA”.

1.2. Rumusan Masalah

Untuk menerapkan algoritma rabin-karp pada program pendeteksi plagiarisme judul, maka rumusan masalah pada penelitian tersebut adalah:

1. Bagaimana cara mengimplementasikan Algoritma Rabin-karp sebagai pembantu pendeteksi plagiarisme sebuah judul dengan menggunakan media website?
2. Bagaimana hasil pengujian Algoritma rabin-karp sebagai pendeteksi plagiarisme dengan menggunakan nilai K-gram yang telah diberikan?

1.3. Batasan Masalah

Suatu batasan masalah diperlukan untuk menghindari pelebaran pembahasan ataupun menyimpang dari pokok pembahasan, agar penelitian terarah dalam pembahasannya sehingga tujuan penelitian tercapai. Beberapa batasan masalahnya adalah :

1. Lingkup pembahasan hanya seputar penggunaan algoritma rabin karp pada perbandingan judul dan tingkat kesamaan pada judul dengan bukti.
2. Penjabaran penggunaan algoritma rabin karp hanya sebatas menampilkan hasil perbandingan berupa berapa persen kesamaan judul yang dimasukkan.

1.4. Maksud Dan tujuan Penelitian

Berdasarkan rumusan masalah yang telah dijabarkan, tujuan penelitian ini adalah untuk menerapkan Algoritma Rabin - Karp untuk mendeteksi berapa persen kesamaan pola pada judul yang telah di inputkan dengan judul lain yang terdapat pada database.

Berdasarkan tujuan penelitian yang hendak dicapai, maka penelitian ini diharapkan mempunyai manfaat dalam ilmu pengetahuan baik secara langsung maupun tidak langsung.

1.5. Manfaat Penelitian

Adapun manfaat penelitian ini adalah dapat menambah wawasan dan pengalaman langsung tentang cara menerapkan Algoritma rabin karp kedalam kasus plagiarisme berbentuk program web, dan bagi pembaca, dapat menambah pengetahuan dan sumbangan pemikiran tentang cara menerapkan Algoritma rabin karp ke dalam program web.

1.6. Metode Penelitian

Metodologi penelitian yang digunakan dalam penelitian ini adalah berjenis kuantitatif, yaitu penelitian yang mengikuti prosedur-prosedur yang telah ditentukan sebelumnya, dan melibatkan penghitungan angka dan matematis dalam implementasi maupun pengujiannya. Metode yang digunakan dalam penelitian ini dijabarkan sebagai berikut:

1.6.1. Metode Studi Literatur

Penulis mencari dan mengumpulkan data penelitian terkait dari berbagai literatur buku Pustaka, internet, jurnal - jurnal ilmiah baik nasional

maupun internasional yang berkaitan dengan masalah yang diteliti sebagai penunjang penelitian.

1.6.2. Metode Pendeteksi Plagiarisme

Metode yang digunakan adalah metode dokumen fingerprint. Dalam metode pendeteksi plagiarisme, terdapat proses sebagai berikut:

1. Pengumpulan Data

Dataset yang digunakan sebagai contoh adalah data sekunder berupa nim, nama dan judul skripsi dari Universitas Sam Ratulangi prodi Informatika.

2. Persiapan Data

Pada tahap ini, data string judul yang telah diambil, direkayasa sedemikian hingga, menjadi baris kata dasar, tahap ini biasa disebut dengan *preprocessing*.

3. Pengolahan Data

Tahap ini merupakan tahap inti dimana data yang telah dipersiapkan akan diolah dengan metode dokumen fingerprint yang menghasilkan bilangan yang disebut dengan nilai hash sehingga dapat dilakukan pengujian.

1.6.3. Metode Pengujian

Metode Yang digunakan untuk Mengukur Berapa Persen Kemiripan kedua judul adalah Sorensen – dice Coefficient. Metode ini digunakan untuk mengukur kesamaan diantara dua set data.

1.7. Sistematika Penulisan

Untuk memahami lebih jelas penelitian ini, maka materi yang tertera pada skripsi ini dikelompokkan menjadi beberapa bab dengan sistematik penyampaian sebagai berikut :

BAB 1 PENDAHULUAN

Bab ini menjelaskan tentang latar belakang masalah, rumusan masalah yang akan dihadapi, batasan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan.

BAB 2 LANDASAN TEORI

Bab ini menjelaskan teori yang berupa pengertian, definisi yang diambil dari kutipan buku, jurnal, ataupun literatur lain yang berkaitan dengan penyusunan laporan skripsi..

BAB 3 METODE PENELITIAN

Bab ini menjelaskan alat dan bahan yang digunakan untuk melakukan penelitian, metode yang digunakan oleh peneliti, alur penelitian secara detail dari awal hingga akhir serta dijelaskan rancangan dan penerapan sistem yang diusulkan.

BAB 4 HASIL DAN PEMBAHASAN

Pada bab ini, serta pembahasan dan hasil dari bab sebelumnya, di jelaskan dengan menerapkan konsep setelah adanya sistem yang telah di usulkan serta dilakukan pengujian terhadap sistem yang telah dibuat.

BAB 5 PENUTUP

Bab ini berisi kesimpulan dan saran yang berkaitan dengan analisis berdasarkan hal yang telah diuraikan pada bab sebelumnya.