

BAB I

PENDAHULUAN

1.1 Latar Belakang

Plagiarisme didefinisikan sebagai mengambil atau mencuri beberapa karya dan mengklaim menjadi miliknya sendiri. Hal ini dapat dilakukan dengan sengaja atau tidak disengaja [1]. plagiarisme terjadi ketika satu penulis mencoba untuk menganggap karya penulis lain sebagai miliknya. Plagiarisme juga terjadi ketika seorang penulis mereferensikan karya orang lain dalam tulisan mereka sendiri dan tidak memberikan kredit yang tepat kepada penulis yang karyanya mereka rujuk.

Konsekuensi dari plagiarisme telah banyak dilaporkan di dunia akademis. Kehilangan kemampuan untuk mempublikasikan kemungkinan besar berarti akhir dari posisi akademis. Akibat hukum dari plagiarisme bisa sangat serius. Seseorang tidak dapat menggunakan materi orang lain tanpa kutipan dan referensi. Mereka yang menulis untuk menjadikan profesi seperti jurnalis atau penulis, sangat rentan terhadap masalah plagiarisme. Konsekuensi dari plagiarisme sangat luas dan tidak ada yang kebal. Sebelum mencoba proyek penulisan apa pun, pelajari dulu tentang plagiarisme. Cari tahu apa yang merupakan plagiarisme dan bagaimana menghindarinya.

Rabin-Karp adalah algoritma yang bagus untuk satu alasan sederhana dapat digunakan untuk mencocokkan dengan beberapa pola. Ini membuatnya sempurna untuk mendeteksi plagiarisme bahkan untuk frasa yang lebih besar. Algoritma rabin karp dengan naive bayes bisa dikatakan sama. Tetapi tidak seperti algoritma Naive, algoritma Rabin Karp mencocokkan nilai hash dari pola dengan nilai hash dari substring teks saat ini, dan jika nilai hash cocok maka rabin karp akan memulai mencocokkan karakter individu. Sedangkan naive bayes kelemahan terbesarnya adalah persyaratan prediktor untuk independen mendeteksi plagiat.

Harapan dari penelitian ini adalah dengan algoritma rabin karp, nantinya bisa digunakan untuk mendeteksi kesamaan teks pada dokumen. Mengetahui seberapa akurat rabin karp untuk mendeteksi dan menggunakan rabin karp untuk membuat aplikasi detector plagiat. Dengan detector plagiat yang tercipta nantinya, maka setidaknya bisa mencegah seseorang, untuk membuat suatu karya ilmiah terhindar dari plagiarisme.

Plagiarisme sangat tidak menguntungkan di dunia pendidikan, dikarenakan bisa membuat mahasiswa tidak bisa lulus tepat waktu. Maka sebaiknya hindari dengan mempelajari terlebih dahulu apakah teks yang kita ketik mengandung kesamaan dengan teks orang lain.

1.2 Rumusan Masalah

Dari latar belakang di atas dapat ditentukan rumusan masalahnya sebagai berikut: berapa akurasi persentase yang didapatkan untuk mendeteksi plagiat dengan rabin karp?

1.3 Batasan Masalah

Dalam penelitian ini, Adapun batasan masalah yang di tentukan dari rumusan masalah sebagai berikut:

1. Penelitian ini hanya mendeteksi kesamaan teks yang ada pada dokumen
2. Dokumen yang di deteksi menggunakan nilai k-gram 5
3. dokumen yang dapat di deteksi hanya bagian abstrak
4. dokumen yang di deteksi sudah menghilangkan karakter-karakter yang tidak perlu

1.4 Tujuan penelitian

Beberapa tujuan yang ingin dicapai penelitian ini sebagai berikut:

1. Mendeteksi dokumen apakah terdapat kesamaan teks menggunakan rabin karp
2. mengecek berapa persentase algoritma rabin karp jika digunakan untuk mendeteksi kesamaan teks

1.5 Manfaat Penelitian

Mengetahui bahwa algoritma rabin karp cocok untuk mendeteksi dokumen yang memiliki nilai kesamaan teks, sehingga nantinya bisa menciptakan system pendeteksi plagiat dan dapat digunakan untuk membuat program atau aplikasi untuk mendeteksi plagiat lebih sempurna di kemudian hari