

Abstrak

Plagiarisme merupakan tindakan menjiplak karya seseorang dan mengakuinya sebagai karya sendiri. Plagiarisme terhadap dokumen susah untuk dihindari. Oleh karena itu detektor otomatis sangat membantu untuk menemukan praktek-praktek plagiarisme pada dokumen seperti MOSS, TESSY, JPlag, COpyCatch dan lain-lain.

Untuk melakukan deteksi plagiarisme dokumen pada intinya dengan melakukan metode yang terdiri dari dua proses sekuensial yaitu mengubah dokumen menjadi representasi linier yang disebut token string, kemudian membandingkan setiap token string satu sama lain. Algoritma yang digunakan dalam skripsi ini adalah Running Karp-Rabin Greedy String Tiling (RKR-GST). Algoritma ini mampu menemukan bagian-bagian yang identik pada dua buah string tanpa terpengaruh urutan maupun posisi substring. Untuk setiap pembandingan dua buah dokumen, nilai similaritas dihitung dan bagian-bagian dokumen yang diduga merupakan hasil praktek plagiarism ditandai.

Aplikasi Algoritma RKR-GST ini dibangun berdasarkan prinsip diatas dengan menggunakan bahasa pemrograman java. Aplikasi ini mampu menemukan dugaan-dugaan praktek plagiarism dan telah diuji menggunakan dokumen hasil pengumpulan dari situs elektronik dan menggunakan dokumen latihan yang telah dimodifikasi.

Untuk pengujian, sebaiknya nilai similaritas tidak menjadi keputusan final untuk menentukan kasus-kasus plagiarism karena ada kemungkinan hasil deteksi salah. Oleh karena itu, untuk dugaan –dugaan plagiarism yang meragukan hendaknya penguji mencermati isi dokumen pasangan tersebut sebelum mengambil keputusan.

Kata kunci : plagiarisme dokumen, *preprocessing*, RKR-GST, *string matching*, *scanpattern*.