Abstrak

Plagiarisme merupakan salah satu tindak pidana yang merugikan orang lain karena plagiarisme berarti mencuri hak cipta orang lain. Plagiarisme terjadi karena rendahnya kemampuan seseorang untuk berinovasi menciptakan suatu karya yang *original*. Tindakan plagiarisme ini sering terjadi di dunia pendidikan salah satunya di perkuliahan. Kasus plagiarisme yang sering terjadi diantaranya adalah plagiat penulisan laporan, tugas, makalah ataupun skripsi mahasiswa. Akan sangat sulit mengukur kesamaan dokumen hanya dengan membaca atau mengukurnya secara manual, ditambah jika yang harus dibandingkan lebih dari dua dokumen. Oleh karena itu, dibutuhkan suatu sistem yang dapat melakukan deteksi plagiarisme*(text reuse detection)* secara otomatis dan akurat. Dalam tugas akhir ini saya telah membuat suatu sistem deteksi plagiarisme menggunakan metode *Text Alignment* untuk deteksi plagiarismenya, dengan melakukan *parsing N-gram* pada setiap dokumen inputnya. Kemudian dilakukan pencarian *string* dengan menggunakan algoritma *Rabin Karp*. Dalam membangun *Text Alignment* dibutuhkan 3 tahapan, yaitu *Seeding*, *Extension* dan *Filtering*. Penggunaan metode *Text Alignment* dalam mendeteksi plagiarisme dapat menghasilkan nilai *Micro Average F1-Score* 82.20% untuk evaluasi level kasus, dan nilai *Micro Average F1-Score* 49.64% untuk evaluasi level karakter. Sistem ini meningkati peringkat 17 untuk *Precision* dan peringkat 16 untuk *Recall* pada *Overview of the 5th International Competition on Plagiarism Detection PAN 2013*.

Kata kunci: text alignment, plagiarism detection, text reuse detection, similarity, seeding, extension, filtering.