

Abstrak

Dalam penelitian ini dilakukan identifikasi plagiarisme antara dua buah dokumen(.txt). Metode yang digunakan dalam penelitian ini yaitu menggunakan *Text Alignment* dan kemudian diolah dengan karakter *N-Gram* dan Algoritma *Winnowing*. Penelitian ini mengolah dua buah dokumen yang berupa *Source Document* dan *Suspicious Document*, kemudian melakukan perhitungan menggunakan *Confusion Matrix* untuk mendapatkan hasil *Accuracy*, *Recall*, *Precision*, dan *F1-Score*. Algoritma *Winnowing* sendiri memiliki keunggulan diantara algoritma dokumen *Fingerprint* yang lain seperti algoritma *Manber* dan algoritma *Rabin-karp*.

Plagiarisme menjadi salah satu permasalahan yang sangat sering ditemukan. Untuk saat ini masih sedikit yang menggunakan teknologi untuk memecahkan permasalahan ini dan lebih banyak hanya dengan *Visual*. Hasil dari penelitian ini berupa *confusion matrix*, dimana dilakukan beberapa skenario yang akan memperlihatkan perfomansi disetiap skenario yang dibentuk. Dari penelitian yang dilakukan, diperoleh hasil penilitian untuk skenario 1 *F1-score* sebesar 0.4624 (46.24%) dan skenario 2 *F1-score* sebesar 0.4719 (47.19%).

Kata kunci : *plagiarism, winnowing, text alignment, n-gram, confusion matrix*