

Abstrak

Maraknya kasus plagiarisme pada saat sekarang ini menjadikan perlu adanya sebuah sistem yang dapat mendeteksi tindakan plagiat. Untuk mengetahui kemiripan satu dokumen dengan dokumen lain yang ada pada *database* maka harus dilakukan pencocokan antara dokumen *query* dengan semua dokumen pada *database*. Proses pencocokan dua dokumen yang memerlukan waktu yang lama menjadi suatu masalah. Semakin banyak jumlah dokumen pada *database* menjadikan waktu komputasi menjadi sangat lama. Untuk mengefisienkan waktu pemrosesan ini diajukan sebuah sistem pendeteksian indikasi plagiarisme dengan menggunakan proses *indexing* dengan struktur data *2-3 tree*. *Indexing* bertujuan untuk mengeliminasi dokumen pada *database* yang benar-benar tidak ada kaitannya dengan dokumen *query*, sehingga proses pencocokan hanya dilakukan pada dokumen yang memiliki keterkaitan saja dengan dokumen *query*. Hasil dari penelitian ini menunjukkan bahwa untuk melakukan pemrosesan satu *query* dan 10000 dokumen *database* membutuhkan waktu 59 detik dengan *indexing* dan 134 detik tanpa *indexing*. Adapun nilai *f-measure*, rata-rata nilai *precision* dan *recall*, yang didapat adalah 0,7387 untuk *indexing* dengan *threshold* eliminasi 0,15 dan 0,000428 tanpa *indexing*.

Kata kunci: plagiarisme, *indexing*, *2-3 tree*, dokumen, *fingerprint*, *query*