ABSTRAK

IMPLEMENTASI TF-IDF UNTUK PENCARIAN DOKUMEN

(Studi kasus di PT. eBdesk Indonesia)

Oleh

Nugraha Iman Santosa
10109797

Mesin pencarian dokumen adalah program komputer yang dirancang untuk melakukan pencarian atas dokumen dalam koleksi dokumen. Hasil pencarian umumnya ditampilkan dalam bentuk daftar yang seringkali diurutkan menurut tingkat akurasi kesesuaian dokumen dengan kebutuhan pengguna. Pada skripsi kali ini akan mengimplementasikan suatu aplikasi pencarian dokumen di dalam aplikasi eBdesk Collaboration. Dimana pencarian dokumen dilakukan terhadap dokumen-dokumen yang dimasukkan pengguna pada aplikasi eBdesk Collaboration.

Ada 2 proses yang dilakukan pada aplikasi yang dibuat, antara lain : proses penentuan indeks dan proses pencarian. Proses penentuan indeks yaitu proses mengekstrak dokumen yang di-*upload*, melakukan *pre-processing* kemudian menyimpan hasilnya ke dalam *database*. Proses selanjutnya adalah pencarian, yaitu mencari dokumen dari database berdasarkan kata kunci dari pengguna, kemudian mengurutkan dokumen hasil pencarian dengan menggunakan metode tf-idf, sehingga dihasilkan daftar dokumen yang telah terurut.

Metode tf-idf adalah metode yang umum digunakan untuk mesin pencarian karena efisien dan sederhana. Pembobotan dokumen menggunakan tf-idf menunjukkan bahwa deskripsi terbaik dari dokumen terjadi ketika kata yang banyak muncul dalam dokumen tersebut dan sangat sedikit muncul pada dokumen yang lain.

Kata Kunci: mesin pencarian, penentuan indeks, *pre-processing*, tf-idf.