ABSTRACT

TF-IDF IMPLEMENTATION FOR DOCUMENTS SEARCHING
(A case study in PT. eBdesk Indonesia)

by

Nugraha Iman Santosa
10109797

Document search engine is a computer program designed to search for documents in document collections. Search results are generally displayed in the form of lists that are often sorted by the accuracy of the document conformance to user needs. In this thesis will implement an application to search documents in the application eBdesk Collaboration. Where the document search conducted on the documents that user uploaded on the eBdesk Collaboration application.

There are two processes are performed on an application made, among others: the process of indexing and the searching process. The process of indexing is a process to extract the uploaded document, perform pre-processing and then save the results into the database. The next process is searching, which is seeking documents from database based on keyword from user, than rank the search result document using tf-idf method, so the result is list of documents that have been ranked.

Tf-idf method is a method commonly used for search engines because of its efficiency and its simplicity. Weighted using the tf-idf document shows that the best description of the document occurs when a word often appears in documents and very rarely appears on other documents.

Keywords: searching machine, indexing, pre-processing, tf-idf.