

Abstract

Information Retrieval (IR) is a part of computer science-related information from the documents that is based on the content and context of the documents themselves. Information Retrieval in the process can be described as a process for the collection of documents Relevant documents through search queries input by user.

Test parameters to assess the relevance of a document that is used in this Final Project is Precision, recall and IAP. Precision is a parameter to calculate the value-level accuracy between a query with the document collection. Precision is the result of the comparison between the document with all relevant documents retrieved successfully by the system. Recall is a parameter to calculate the value of the level of completeness of a query with the document collection. Recall is the result of the comparison between the relevant documents with the relevant documents that exist in the entire document collection. IAP keterurutan calculate the value of relevant documents in the system. IAP values obtained from the calculation of the value of Precision and recall.

In the IR there is a model for the value similiarity documents relevant to the query input by by the user. One model is the Vector Space Model. There are 3 methods weightining in Vector Space Model, the Term Frequency (TF), inverse Document Frequency (IDF) and Term Frequency-inverse Document Frequency (TF-IDF). From the test results obtained with the parameters that IAP TF-IDF method is superior compared with the other two methods. This shows that the method weightining TF-IDF is better than the other 2 methods, namely TF and IDF.

Keywords: *Information Retrieval, Information Retrieval System, Vector Space Model, precision,recall, IAP , TF, IDF, TF-IDF , document collection and query.*