

ABSTRACT

A search engine is required to give a precise search results and more *Relevant* to the user desires. The concept of search engines is to use information retrieval. In information retrieval, there are two types of documents, namely the free text (unstructured document) and fielded text (structured document). HTML document is one of the fielded text. In the search for documents in HTML form, must be considered a factor of importance of each part of the document. The importance of documents based on meta tag title, keywords, body, and H1 and *Pagerank* using Okapi method (*BM25F*) to generate the level of performance required parameters.

BM25F method implemented in the scope of the document weigh (weighting), while the *Pagerank* method is implemented on the document node weight (ranking). To combine both methods, each method produces values which will then be multiplied to determine the similarity score of each document is calculated. By applying both methods are then combined, will produce a level of performance possible from the parameter (*Precision*, *Recall* and Interpolated Average *Precision*) is better than just a stand-alone methods such as only using *BM25F* method.

The test results on the test scenario Top-N changes, it can be concluded that the results achieved stable system performance in some queries, this occurs because the query is a document that has a high level of similarity scores and *Pagerank* value that can be said to be large or high-level document kepopularan. Combining these methods greatly affect the *Pagerank* value of a document, because if the *Pagerank* value of the resulting similarity score will also be higher than other documents. However, combining methods and *BM25F Pagerank* method has drawbacks in parameter *Recall* (completeness) produced will be decreased to <50% instead of using a method *BM25F* only. As with the parameters and *IAP (INTERPOLATED AVERAGE PRECISION)* *Precision* better than *BM25F* method that can achieve a percentage of 100%.

Keywords: Information Retrieval, *BM25F* method, Method *Pagerank*