

ABSTRAK

Ketersediaan informasi dalam jumlah besar di satu sisi merupakan berkah yang patut disyukuri pada zaman ini, karena dengannya kehidupan manusia akan semakin mudah. Namun di sisi lain, melimpahnya informasi ini melahirkan permasalahan baru bagi manusia. Yaitu, perihal bagaimana mendapatkan informasi yang dibutuhkan secara tepat dan akurat diantara kumpulan berbagai informasi. Untuk itu dikembangkanlah cara untuk mengatasi permasalahan ini yaitu sistem temu kembali informasi atau *information retrieval*. Suatu sistem temu kembali informasi yang baik memiliki tingkat relevansi yang bisa diterima oleh pengguna. Untuk dapat menghasilkan nilai relevansi yang tinggi, maka salah satu caranya, sistem ini perlu menerapkan metode perangkungan yang baik dan teruji. Kemudian yang menjadi pertanyaan, bagaimana menentukan suatu kinerja metode perangkungan. Kinerja metode perangkungan ditentukan oleh relevansinya yang diukur dengan parameter precision dan recall. *Latent semantic analysis* pada tugas akhir ini akan dikombinasikan dengan *cosine similarity*, sehingga untuk mengukur kinerjanya perlu diimplementasikan ke dalam perangkat lunak untuk kemudian diuji parameternya.

Dalam suatu pengujian diperlukan metode lain sebagai pembanding untuk mengukur kinerja *latent semantic analysis* yang dikombinasikan dengan *cosine similarity*, maka dipilihlah *vector space model* sebagai pembanding.

Hasil pengujian pada tugas akhir ini menunjukkan bahwa *latent semantic analysis* memiliki nilai similaritas yang lebih baik dari *vector space model* dengan nilai precision 81,5% dan nilai recall sebesar 80,2% sedangkan nilai f-measurenya 80,8

Kata kunci: *Latent Semantic Analysis, Cosine Similarity, Vector space model*

ABSTRACT

The availability of large amounts of information on one side is a blessing to be thankful for at this time, because of her human life will be easier. But on the other hand, the abundance of information have spawned new problems for humans. Namely, regarding how to obtain the needed information appropriately and accurately among a collection of various information. For that it is developing a way to overcome this problem is information retrieval. An information retrieval system that either have a level of relevance that can be received by users. To be able to produce high relevancy value, then one way, these systems need to apply the method of ranking the good and tested. Then the question becomes, how to determine a method of ranking the performance. Performance is determined by the method of ranking the relevance measured with precision and recall parameters. Latent semantic analysis in this final will be combined with the cosine similarity, so as to measure its performance needs to be implemented into the software and then tested parameters.

In a test other methods are needed for comparison to measure the performance of latent semantic analysis combined with cosine similarity, the chosen vector space model as a comparison.

The test results in this thesis show that latent semantic analysis has a value of similarity better than vector space model with a value of 81.5% precision and recall value amounted to 80.2% while the value of f-measure of 80.8

Keyword: *Information Retrieval, Latent Semantic Analysis, Cosine Similarity.*

