

ABSTRAKSI

Salah satu cara untuk membantu memahami sebuah konsep adalah dengan mencari hubungan antara konsep tersebut dengan konsep yang lainnya. *Latent semantic analysis* adalah salah satu penerapan dari hal tersebut. Dimana LSA dapat mencari nilai kesamaan antar konsep. *Latent semantic analysis* (LSA) adalah suatu cara untuk mengekstraksi tulisan dan membuat representasi statistik dari tulisan tersebut dengan pendekatan dimensi rendah (*low rank approximation*), dimana digunakan dimensi yang kecil untuk mengetahui keseluruhan makna dokumen .

Tugas Akhir ini menguji efektifitas penggunaan metode *latent semantic analysis* yang merupakan proses analisis semantik pada *short text* seperti al-Quran untuk pencarian nilai kesamaan makna ayatnya sesuai makna yang terkandung dari tiap ayat. Selain itu, mengingat LSA bersifat dinamis, dimana tidak ada cara yang pasti untuk menentukan berapa dimensi yang harus digunakan untuk setiap kasus, maka perlu diketahui bagaimana pengaruh pemilihan dimensi pada nilai *similarity*.

Berdasarkan hasil pengujian, *term-document matrix* tanpa LSA memiliki akurasi yang tinggi yaitu 85% tetapi berdasarkan pengukuran F-measure, efektifitasnya sangat rendah yaitu 0% karena tidak bisa mendeteksi ayat yang maknanya serupa. Dibandingkan dengan pengukuran dengan LSA walaupun akurasi tertinggi dari salah satu variabel 71% tapi nilai F-measure mencapai 40% yang dimiliki oleh variabel LSA yang lain. Selain itu dimensi yang digunakan berpengaruh pada hasil. Jika terlalu rendah maka nilai kesamaan akan terlalu seragam, bias dan hasilnya tidak akan optimal.

Kata Kunci: *latent semantic analysis, similarity, al-Quran, matriks, singular value decomposition, cosine similarity*