

Rancang Bangun Aplikasi Search Engine Dan Sistem Pendeteksi Plagiarisme Menggunakan Metode LSA-SOM

**Aris Prama Aulia
(08.04.111.00002)**

Pembimbing 1 : Hermawan ST, M.Kom

Pembimbing 2 : Dr. Indah Agustien S, S.Kom, M.Kom

ABSTRAK

Plagiarisme adalah mencuri gagasan, kata-kata, kalimat atau hasil penelitian orang lain dan menyajikannya seolah-olah sebagai karya sendiri. Metode LSA merupakan salah satu metode dari beberapa metode untuk pendeteksian plagiarisme. Metode *Latent Semantic Analysis* (LSA) adalah sebuah teori dan metode untuk menggali dan mempresentasikan konteks yang digunakan sebagai sebuah arti kata dengan memanfaatkan komputasi statistik untuk sejumlah *corpus* yang besar. Penggunaan LSA konvensional bersifat sekuensial dikarenakan masih belum adanya klusterisasi sehingga proses analisa akan dilakukan secara global pada keseluruhan dokumen yang tentunya memerlukan waktu komputasi yang besar. Untuk itu dapat digunakan teknik klasifikasi tak terawasi seperti penggunaan *Self Organizing Map* (SOM) dan pengujian similaritas secara global hanya dilakukan pada centroid masing-masing kluster untuk kemudian dilakukan pengujian similaritas secara lokal pada kluster yang terpilih yang menjadi bagian dari kluster tersebut. Pada hasil penelitian LSA-SOM memiliki rata-rata *precision* 94.4% dan *recall* 64.04%.

Kata kunci : *Cosine Similarity, Latent Semantic Analysis, plagiarisme Self Organizing Map, Singular Value Decomposition*