

Abstrak

Adanya pengelompokan suatu dokumen sangat diperlukan untuk berbagai bidang teknologi. Perkembangan data pun kian meningkat dan kebutuhan akan pemrosesan data kelompok yang cepat dalam keadaan *real time* pun semakin dikembangkan.

Cover Coefficient Clustering Method (C2ICM) adalah salah satu algoritma pengelompokan atau *clustering* dokumen yang menggunakan model probalistik, kesamaan *term*, dan *seed* dokumen sebagai poin penting dalam menentukan inisialisasi awal dari pembentukan *cluster*-nya, dimana dokumen yang dikelompokkan ini telah diproses terlebih dahulu agar bisa diolah menggunakan algoritma *C2ICM* dan memenuhi kebutuhan akan dinamika perkembangan data yang selalu terjadi *update real time*. Pemrosesan ini disebut juga *preprocessing* atau *indexing* dokumen. Dalam *Indexing*, ada beberapa tahap yang biasa dilakukan, salah satunya *stemming*. Pada Tugas Akhir ini, Algoritma Jelita Asian digunakan pada tahap *stemming* kata dengan dilanjutkan penggunaan Algoritma *C2ICM* dalam pengclusteran dokumen abstrak TA mahasiswa ITTELKOM.

Pengujian yang dilakukan dalam Tugas Akhir ini yaitu untuk meng-*cluster*-kan dokumen dengan Algoritma *C2ICM*, menganalisa konsep *Cover Coefficient C2ICM* dan tahapan Algoritma *C2ICM* serta menganalisa hasil kualitas *cluster* yang dihasilkan oleh perpaduan Jelita Asian dengan *C2ICM* menggunakan nilai *Silhouette Coefficient*. Dan didapatkan bahwa kualitas *cluster* yang dihasilkan termasuk dalam kualitas yang lemah berdasarkan nilai rata-rata *silhouette* yang diperoleh.

Kata kunci : *C2ICM, Jelita Asian, SilhouetteCoefficient, Clustering, Indexing*