

Abstraksi

Dalam pembuatan tugas akhir ini digunakan metode *Cover Coefficient-Based Incremental Clustering Methodology* (C2ICM) yang dilakukan pengujian terhadap koleksi dokumen SMART, yaitu data ADI dan CISI. Koleksi dokumen tersebut dilakukan pemrosesan sedemikian sehingga terbentuklah *cluster-cluster* yang memiliki kualitas yang beragam. Adapun tahapan-tahapan yang dilakukan untuk membentuk *cluster-cluster* tersebut meliputi *preprocessing*, pembentukan D-Matrix, perhitungan C-Matrix, perhitungan jumlah *cluster*, pemilihan *seed* dokumen, dan pengelompokkan dokumen *nonseed* ke dalam *seed* dokumen terpilih. Ketika terdapat dokumen baru memasuki *database*, maka proses selanjutnya yang akan dilakukan adalah *incremental maintenance*. Jika *seed* dokumen terdahulu terpilih kembali menjadi *seed*, maka *cluster* yang telah terbentuk dari proses sebelumnya akan tetap digunakan, namun bila *seed* dokumen terdahulu tidak terpilih kembali, maka *cluster* yang telah terbentuk akan dihapus, dan akan dibentuk *cluster* baru sesuai dengan *seed* dokumen yang baru. Setelah terbentuknya *cluster* untuk semua *query*, maka akan dihitung kualitas *cluster* dengan menggunakan *Silhouette Coefficient* (SC) dan waktu eksekusi pembentukan *cluster*. Dari hasil percobaan, kualitas *cluster* yang terbentuk dari penggunaan algoritma C2ICM memiliki karakteristik/kualitas yang beragam. Sedangkan waktu eksekusi yang dihasilkan menunjukkan waktu yang linier terhadap pertambahan jumlah dokumen *hitlist*.

Kata kunci: *incremental, cluster, seed, cover coefficient, silhouette coefficient, kualitas cluster*