

ABSTRAK

Clustering merupakan salah satu metode dalam *data mining* yang sering digunakan dalam penelitian, yang mempunyai tujuan untuk mengelompokkan data tanpa adanya label kelas atau yang sering disebut *unsupervised*. Dalam tugas akhir ini akan membahas tentang algoritma *clustering* yang berusaha mengelompokkan data mahasiswa ke dalam sejumlah *cluster*.

Algoritma yang digunakan adalah gabungan 2 (dua) algoritma *clustering* yaitu ISMC (*Improved Split and Merge Classification*) dan FCM (*Fuzzy C-Means*). Kedua algoritma tersebut memiliki peran yang berbeda. Untuk ISMC digunakan untuk menentukan jumlah *cluster*, sedangkan FCM digunakan untuk menentukan anggota tiap *cluster*.

Analisa performansi algoritma menggunakan 2 (dua) parameter yaitu *between cluster scatter matrix* S_b dan *within cluster scatter matrix* S_w . Dari hasil pengujian performansi dengan menggunakan 2 (dua) set data yaitu data bunga iris dan data mahasiswa diperoleh bahwa S_w algoritma gabungan mengalami perbaikan. Untuk data set bunga iris dapat mencapai 81%, sedangkan data set mahasiswa mencapai 99% lebih baik dibanding dengan algoritma ISMC. Hal tersebut dapat disimpulkan bahwa algoritma gabungan mampu menghasilkan *cluster* yang lebih homogen dibanding dengan ISMC.

Kata Kunci : *Data mining*, Algoritma *clustering*, *between cluster scatter matrix*, *within cluster scatter matrix*, *data bunga iris*, *data mahasiswa*