

Abstraksi

Clustering dokumen berita diperlukan untuk dapat mengelola informasi menjadi sebuah pengetahuan yang berguna. Salah satu metode yang dapat digunakan untuk klasterisasi dokumen berita ini adalah metode *Fuzzy Hierarchical Agglomerative Clustering* (FHAC). Metode FHAC mempunyai kemampuan untuk menemukan, menganalisa, dan menggabungkan daerah data dengan cepat. Dengan menggunakan fungsi evaluasi *fuzzy*, metode FHAC dapat menemukan skema *clustering* yang paling optimal. Kualitas *cluster* yang dihasilkan bergantung dengan nilai beberapa parameter inputan, yaitu A (*threshold merging cluster*) dan K (jumlah tetangga). Untuk data yang ketidakmiripannya kecil/ jarak antar titik kecil/ jumlah label dokumen sedikit lebih tepat menggunakan nilai K yang besar dibandingkan dengan penggunaan nilai K yang kecil, begitu pula sebaliknya pada data dengan kemiripan yang besar/ jarak antar titik besar/ jumlah label dokumen banyak menggunakan nilai K yang kecil. Setiap *cluster* yang telah terbentuk diberi label yang paling merepresentasikan *cluster*, yaitu dengan melihat label dominan dalam *cluster*. Kualitas *cluster* dan akurasi pelabelan *cluster* diukur dengan *silhouette coefficient* dan *precision*.

Kata kunci : Klasterisasi, FHAC, Pelabelan, Silhouette Coefficient, Precision