

Abstrak

Jumlah dokumen sangat banyak dan harus ditangani memerlukan sistem pengorganisasian secara otomatis. Salah satu teknik dalam pengorganisasian dokumen adalah clustering. Pada tugas akhir dibandingkan performansi dari dua algoritma *clustering* yaitu *K-means* dan *Suffix Tree Clustering* dalam mengelompokkan berita dalam bahasa Indonesia. Pengukuran performansi yang dimaksud adalah *recall*, *precision*, *F-measure* dan *variance*. Dengan parameter tersebut akan terlihat algoritma terbaik dalam mengelompokkan berita dalam bahasa Indonesia. Algoritma *K-means* memiliki kerapatan cluster yang lebih baik dibandingkan dengan *Suffix Tree Clustering*, sedangkan *Suffix Tree Clustering* memiliki nilai *F-measure* yang lebih tinggi dibandingkan *Kmeans*.

Kata kunci: *clustering, K-means, Suffix Tree Clustering recall, precision, F-measure dan variance*