

Abstrak

Saat ini perkembangan dunia teknologi komunikasi dan informasi semakin pesat. Begitu juga dengan jumlah dokumen berita khususnya berita digital yang ada pada media online. Hal ini menyebabkan semakin sulitnya untuk melakukan pencarian terhadap suatu topik berita. *Clustering* merupakan salah satu metode data mining yang bersifat *unsupervised learning* untuk mengelompokkan dokumen berdasarkan kemiripannya. Untuk melakukan pengelompokan tersebut, digunakan salah satu algoritma *clustering* yaitu *Partitioning Around Medoid (PAM)*. *Partitioning Around Medoid* merupakan suatu algoritma *clustering* yang berusaha melakukan pengelompokan suatu dataset dengan mencari terlebih dahulu sejumlah titik yang merepresentasikan suatu *cluster (medoid)*. Setelah mendapatkan k *medoid* dokumen pada dataset dikelompokkan kedalam cluster yang memiliki jarak ke *medoid* terdekat. Adapun metode pendekatan yang digunakan untuk menghitung jarak antar dokumen adalah *euclidean distance method*.

Dalam Skripsi ini kualitas *cluster* yang dihasilkan diukur dengan dua parameter yaitu *cohesion* dan *separation*. Nilai *cohesion* menunjukkan jarak antar objek pada sebuah *cluster* sedangkan *separation* adalah jarak antar objek pada *cluster* yang berbeda.

Kata kunci: *partitioning around medoid, medoid, unsupervised learning, euclidean distance method, cohesion, separation.*