Abstract

Clustering is defined as non supervised classification of patterns of data into groups is called a cluster. The object is for the case of distribution (people, objects, events, etc.) into groups, such that the degree level will be strong connectedness between members of the same cluster and weak between members of different clusters. Clustering is performed on documents called Document Clustering.

In this Final Project implemented a clustering algorithm that works in an incremental document to handle the increasing cases of data with case studies TA document. One of document clustering algorithms is Chung-McLeod algorithm. Chung-McLeod algorithm differs from other incremental algorithms because it can handle the problem of the order of arrival of documents.

Tests carried out by calculating the accuracy of the document clustering software that uses Chung-McLeod algorithm with the results of clustering are considered a golden standard with the same input data. Accuracy obtained from this test is about 75%. Beside that, also seen influence the order of arrival of data on the results of the cluster software with this algorithm.

Keywords: Document *Clustering*, *incremental*, *Chung-McLeod algorithm,*