ABSTRACT

Spreading the news today increasingly widespread since the more rapid development of the Internet world. These developments make the news spread increasingly diverse and vast numbers. News readers will find it hard to obtain the desired news if the news is not clustered well. And if it should be classified manually takes a very long time. Therefore, *Clustering* becomes a solution to resolve the issue. *Clustering* will categorize news documents based on similarity of the document.

Method of single linkage hierarchical *Clustering* is a method of grouping. Single linkage method of classifying documents based on the shortest distance between the documents. Group variation of the data as a group until all the data merged into a single group. Computing single linkage is computationally expensive and complex. While the method of k-means *Clustering* is a method of grouping partitioned. K-means *Clustering* method to group documents based on the closest distance to the centroid. K-Means *Clustering* is a method that is simple and can be used easily. In certain types of data, K-means *Clustering* segmentation can not provide the data so well that the group formed impure same data.

Test methods used to measure the quality of a cluster is the Silhouette Coefficient and Purity. Based on the test results, Single Linkage method has better performance than the K-means clustering method. For the value of silhouette coefficient, Single Linkage is always better than K-Means. Increase the number of documents made silhouette coefecient single linkage value getss smaller while the K-means sometimes produce negative values. For the value of purity, Single Linkage is always equal to 1 while the K-Means is never worth 1. This means that the single linkage always produce the same document, while K-means is still mixed with other documents.

Keywords: Clustering, HAC, Partitioned, Single Linkage, K-Means, Silhouettte Coefficient, danpurity.