Abstract

In the current era, computing and information has been growing fast. Especially in the field of data processing . Along the time, the created data keeps increasing . Therefore we need a solution so that the data can be processed into useful information . Data mining is a solution to handle the amount of data. Data mining is a tool that has an important role to organize and process the amount of data.

In this research, the author discusses the clustering of data for text data, which uses data taken from the online site news and moved to a notepad with a .txt format. In this study, the authors use the Single Linkage clustering. Single linkage clustering method is part of agglomerative hierarchical clustering. Single linkage method of classifying the data based on the data or object nearby. The method used to determine the quality of the cluster is the validity of the similarity oriented.

The results from the system summarize and single linkage clustering is good enough. A percentage of 50% was the best in Summarize, which has a 74.10% accuracy rate. Meanwhile, Single linkage clustering with 12 clusters have better accuracy rate than single linkage clustering with 6 cluster which was 79.8%.

Keywords *: Data mining, text mining, clustering, agglomerative hierarchical clustering, a single linkage, summarize*