

## Abstract

Today information technology is growing rapidly, causing the number of documents online and offline news are also multiplying. Because the number of news documents that there are many, the search for information on certain topics in a collection of news documents tend to be more difficult. This requires grouping of news articles by applying one of the techniques associated with the Text Mining is Clustering.

Clustering is one method of data mining is unsupervised learning to classify documents based on similarity. To perform these groupings, in this Final used one that is divisive clustering algorithm hierarchy Analysis (DIANA). Divisive Algorithm Analysis (DIANA) is top-down clustering where at first all objects are grouped in the same cluster, the object that has the greatest dissimilarity with other objects will be separated into a new cluster, on the other objects would be seen if it has in common with the cluster new, then the object will be grouped into clusters, if the object has no similarity to the cluster with a cluster of new and original then the object will be categorized into a new cluster again, and so on until eventually there is at least one single object in a cluster .

In this Final testing how the clusters are generated and the evaluation of the cluster by performing tests on divisive Coefficient (DC), the relationship between the minimum distance to the F-measure and the relationship between the minimum distance by the number of clusters generated. From the test results that have been conducted where the number of different documents and the smallest number of terms which have the maximum condition on the minimum distance sebsar DC 7.5 which has a value of 0.6812 and a total value of the F-measure of 0.4154. While the input dataset that has the same number of documents with the same number of terms at most, of the conditions at the minimum distance a 22.5 that has a DC value of 0.6323 and a total value of the F-measure of 0.4071.

**Keywords** : Clustering, Divisive Analysis (DIANA), Divisive Coefficient (DC), F-measure, minimum distance.