Abstract

With the advancement of technology, the human ability grow even further in collecting and processing data which are computerized and resulted more data in digital form. Data stored must have various informations. For easy retrieval of information from these data is necessary to group data automatically. Clustering is a process for classifying data into clusters, so that an object in a cluster have high similarity with other objects in the same cluster, but is not similar with objects in other clusters. To perform the clustering process, one of the clustering algorithm used is Jarvis Patrick clustering. Jarvis Patrick clustering is a clustering method based on the similarity between the nearest neighbors. One or more neighbors usually are used to calculate the cluster membership of the object being clustered. The cluster formed depending on the value of parameter J and K. To measure the similarity between the documents before the clustering process, Euclidean distance method is used.

In this final project the clusters quality is measured using silhouette coefficient parameters. Based on experiments conducted that average silhouette coefficient value is 0,220. The quality of the cluster is "no structure", which means becomes practically impossible to find significant cluster centers and to definitely assign the majority of data points.

Keywords: Jarvis Patrick, Euclidean distance, *Silhouette coefficient*, *Clustering*