Abstract

Each institution likes government institution, education institution, non-government institution, and other must be have at least one data for their business. One of them is two dimensional data. It is textual data. Two dimensional data have two basic structure, that is row and column. That data must to process to result a information. One of the data processing is data classification.

Similarity-Driven Cluster Merging is one of data classification method which mayor in similarity between many data in data population or cluster. Similarity-Driven Cluster Merging method can be use for calculate cluster in the data whenever in classified data or non classified data.

In this thesis, Similarity-Driven Cluster Merging method is used for data that haven't classification structure(unsupervised).similarity matrix is used for calculating and saving similarity values in cluster, there for we can determine the similarity of data. Beside of them, it also used for Generalized Objective Function to update the membership value data for a cluster.

From the result testing and result analysis, Similarity-Driven Cluster Merging method resulting high quality cluster with error epsilon at 50 and threshold cluster merging 1.2. High quality cluster was shown with high score function value at 0.914686.

Keyword : *Similarity-Driven Cluster Merging, cluster, merge threshold, Generalized Objective Function.*