Abstract

Clustering is a method of organizing and managing data into classes or cluster that data in a cluster will be in a high equality each other, but it will be different with another in other cluster. In my pilot project, process of clustering, it is used method Fuzzy C-Means. Before the process occurs, first, it does processing data into collection of text, such as words, then it is given a weigth parameter, that's index from how many times each words appears. To evaluate the quality of clustering, it is used a concept validating cluster of XB (Xie-Beni) Index, namely it counts data density in cluster and distance of minimum separity from those cluster.

Keyword: Clustering, Fuzzy C-Means, Text Mining, XB Index