Abstract

Clustering is one of the data mining functionality that is used for grouping the data into a class or cluster. The basic principle is to classify objects klasterisasi in a class that has a great similarity with other objects in the same class (similarity), but it is not similar to objects in other classes (dissimilarity). There are several techniques, among others klasterisasi: Partitioning methods (K-means Clustering), a method Hierarchy (Divisive and Agglomerative Clustering), Density-based methods (DBSCAN), and so on.

In this final task, used partition method with PSO algorithm + K-means (Particle Swarm Optimization + K-means) which is a merger between K-means algorithm and the algorithm of PSO (Particle Swarm Optimization). clustering method using K-means is very much used to classify the data with a high similaritas. However, K-means has a weakness in determining the initial cluster center is done at random so often trapped in local causes and results of the optimal clustering be not optimal. To further optimize the determination of the center of the PSO algorithm is used + K-means. PSO + K-means algorithm which is proposed by Xiaohui Cui in 2005. This algorithm always produces convergence in the global optimum. PSO + K-means was able to avoid local optimum. In this final task will be performed comparative evaluation of the results generated by Clustering method Clustering K-means and PSO with ADVDC (average distance documens cluster centroid) and time performance.

Key : *clustering, particle Swarm Optimization,PSO, K-means, similarity.*