

Abstract

Clustering is one of data mining functionalities which is used to group data into classes or clusters. The basic principle of clustering is to group the object into cluster which has many similarities with other object in the same cluster and many dissimilarities with other object in different cluster. There are several clustering techniques, such as Partition method (K-means Algorithm), Hierarchical method (Divisive and Agglomerative Clustering), Density-Based method (DBSCAN), etc.

This final project is implemented Partition method with FGKA (Fast Genetic K-means Algorithm) which is merged by K-means algorithm and Genetic algorithm. K-Means algorithm is often used to group data that have many similarities. Nevertheless, K-Means has a weakness on determining centroid initial cluster point which is done randomly so that it causes K-Means trapped on local optimal and the result of clustering cannot be optimal.

To make the best of centroid point determining and minimal time, algorithm FGKA can be used. FGKA is algorithm which is developed from GKA algorithm proposed by Yi Lu in 2004. This algorithm always converges to a global optimum eventually. FGKA and GKA can avoid local optimum, but FGKA runs faster than GKA. This final project also compares the result of clustering evaluation from a clustering software with K-Means method.

Keywords : *clustering, GKA, K-Means, FGKA, Fast Genetic K-Means Algorithm*