

Abstract

Many companies may not aware that various data in their database can be processed to help them to develop their business. A particular methodology that can be used is process the data and obtain its benefit is data mining. One methodology that is mostly used in data mining is clustering. Clustering is a classification of various object based on their characteristics. Clustering can support company to identify company's data pattern and to be able to see the tendency of the data.

This final project explains the implementation of k-harmonic means, one of clustering methodology, which is implemented in the customer's billing database. K-harmonic means is a variance of one of the most famous *clustering* methodology, k-means. K-harmonic means has the ability to overcome the weakness of k-means which is incapability to initialize center point. K-means also not able to work in optimum level in bad center point initialization environment. Testing will be done to identify the quality of clustering based on a validation method named silhouette coefficient.

Based on evaluation has been done, K-harmonic means can work in optimum level both in random initialization and bad initialization compared to k-means. In addition, k-harmonic means can be applied in segmentation customer's billing which is the result of clustering have different characteristic each other.

Keywords : *data mining, clustering, k-means, k-harmonic means, silhouette coefficient*