ABSTRACT

Clustering is one of data mining method which is used often on a research that has a purpose to cluster or grouping data without class label which is called unsupervised. In this final task will research about clustering algorithm which is try to group university student's data into some cluster.

The algorithm which is used is merging two clustering algorithm, they are ISMC (Improved Split and Merge Classification) and FCM (Fuzzy C-Means). Those two algorithm have a different character. ISMC is used for determining the amount of cluster, and FCM is used for determining the member of each cluster.

The algorithm performance analysis using two parameters, they are "between cluster scatter matrix Sb" and "within cluster scatter matrix Sw". From the performance testing result with two data set, bunga iris data and university student's data got the merge Sw algorithm is getting better. For bunga iris data set can reach 81%, and university student's data set reaches 99%, better than ISMC algorithm. Those conclude that merge algorithm produce more homogen cluster than ISMC.

Keywords : *Data mining*, *clustering algorithm, between cluster scatter matrix, within cluster scatter matrix, bunga iris data, university student's data*