

Abstract

K-Modes is an extension from K-means clustering algorithm for handling categorical data where modes is used instead of means. Simple K-Modes use simple matching measure to decide similarity value from a cluster. This research will modified k-modes algorithm in deciding similarity value using weighted dissimilarity measure because the value from an attribute really affect a clustering process. The experiment will be tested using real world data sets obtained from the UCI data repository where the type of the data sets is categorical. The performance of the algorithm will be seen from the value of cluster purity from created clusters. From the test showed that the value of cluster purity using k-modes algorithm using weighted dissimilarity measure has the better result from the original k-modes. It shows that object is more accurate allocated in their respective cluster using the new improved k-modes.

Key Word: *clustering, categorical data, weighted dissimilarity measure, cluster purity*