

Abstract

Clustering is one of data mining techniques that splits and classifies data based on similarity from the characteristic of the data. There have been many algorithms that can apply clustering on numerical data. But on categorical data, which is data that categorical where an object holds certain value on specific attribute and each attributes owns set of value, is still few. One of algorithms that can apply clustering on categorical data is Min Min Roughness (MMR). MMR makes use of rough set theory with the number of clusters to be formed as an input parameter, where rough set theory will obtain roughness of each value on certain attribute toward the other ones. From each value roughness on certain attribute, mean roughness will be obtained that will be used to choose attribute and its value as a basis on the splitting of data. From the clusters, will be analyzed on each clusters with purity method and the objects that belong on the cluster.

Keyword : Clustering, MMR, Rough Set Theory