Abstract

A spatial data mining is a process of extracting valid and useful information out of generated data, which recently becomes a highly demanding field due to the huge amount of data collected everyday across various applications domains which by far exceeded human’s ability to analyses, this brought about the development of many data mining tools among which clustering is recognized to be the efficient data mining method that categorized data based on similarity measures, where k-Means is a well-known clustering algorithm used across different application domains. Similarly, k-Means suffer from multiple limitations with its clustering accuracy fully depend on cluster center positioning. In this paper, a density base k-Means cluster centroid initialization algorithm has been proposed to overcome k-Means’s cluster center initialization problem. To prove the accuracy of the proposed algorithm the evaluation test was conducted using two synthetic datasets called Jain and Path base dataset. The clustering accuracy result of the proposed algorithm is compared with that of traditional k-Means algorithm where it proved that the clustering accuracy of the proposed algorithm is better than that of traditional k-Means algorithm.
Density base k-Mean’s Cluster Centroid Initialization Algorithm

References


Index Terms

Computer Science

Algorithms

Keywords

Temporary Matrix(TMAT), cluster center C, Dataset D.