Abstract

Clustering is a division of data into groups of similar objects. Clustering is an unsupervised learning, due to its unknown label class in the search domain. K-means is one of the simplest
unsupervised learning algorithms that solve the well known clustering problem. It has capability to cluster large data. The main idea of K-Means is to define k centroids for each cluster. The K-means algorithm clusters the data with more complexity and the complexity further increases based on the dimensionality and data size. To overcome this we present a novel approach called perimeter K-means (PKM) clustering algorithms, which considers two data points and evaluates the perimeters. From this the two data points are assigned to the nearest cluster center. By this the PKM reduces the overall complexity issues of K-means algorithms. The experimental result on various datasets, with various instances clearly indicates the efficacy of the proposed method. Further cluster quality and stability issues are tested by the proposed PKM.

References


Index Terms

Computer Science

Data Mining

Keywords

Data Mining

Clustering

Similarity

Stability