

## Ranked k-medoids clustering

### Abstract

Cluster analysis is the process of dividing a set of objects into none-overlapping subsets. Each subset is a cluster, such that objects in a cluster are similar to one another and dissimilar to objects in other clusters. One of the most well-known approaches of clustering is partitioning approach involving widely used algorithms like k-means and k-harmonic means. These algorithms suffer from trapping in local optimum; moreover, k-means is sensitive to initialization and outliers. K-medoids algorithms are attempts to reduce the effect of outliers and speed up finding clusters; however, they still have problems with sensitivity to initialization and local optimum. In this paper, we introduce a novel k-medoids algorithm called ranked k-medoids which is not sensitive to initialization and can escape from local optimum. A ranking function derived from similarity among objects is used to find the medoids. In this algorithm the similarity between pairs of objects are computed once and finding the centrally located objects is regardless of the size of dataset. Comparison between our algorithm, k-harmonic means and a new k-medoids algorithm called '*simple and fast k-medoids*' is performed by using four well-known external validation measures over two artificial and five real datasets. The results of our algorithm especially for larger datasets show the superiority of the proposed algorithm over two other algorithms in terms of speed and accuracy.

### Keywords

Partitional Clustering; K-medoids; K-harmonic means; External validation measures;