

A Minimum Spanning Tree-Based Method for Initializing the K-Means Clustering Algorithm

Authors : J. Yang, Y. Ma, X. Zhang, S. Li, Y. Zhang

Abstract : The traditional k-means algorithm has been widely used as a simple and efficient clustering method. However, the algorithm often converges to local minima for the reason that it is sensitive to the initial cluster centers. In this paper, an algorithm for selecting initial cluster centers on the basis of minimum spanning tree (MST) is presented. The set of vertices in MST with same degree are regarded as a whole which is used to find the skeleton data points. Furthermore, a distance measure between the skeleton data points with consideration of degree and Euclidean distance is presented. Finally, MST-based initialization method for the k-means algorithm is presented, and the corresponding time complexity is analyzed as well. The presented algorithm is tested on five data sets from the UCI Machine Learning Repository. The experimental results illustrate the effectiveness of the presented algorithm compared to three existing initialization methods.

Keywords : degree, initial cluster center, k-means, minimum spanning tree

Conference Title : ICIST 2017 : International Conference on Information Science and Technology

Conference Location : Singapore, Singapore

Conference Dates : January 08-09, 2017