World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:11, No:06, 2017

A Computational Cost-Effective Clustering Algorithm in Multidimensional Space Using the Manhattan Metric: Application to the Global Terrorism Database

Authors : Semeh Ben Salem, Sami Naouali, Moetez Sallami

Abstract : The increasing amount of collected data has limited the performance of the current analyzing algorithms. Thus, developing new cost-effective algorithms in terms of complexity, scalability, and accuracy raised significant interests. In this paper, a modified effective k-means based algorithm is developed and experimented. The new algorithm aims to reduce the computational load without significantly affecting the quality of the clusterings. The algorithm uses the City Block distance and a new stop criterion to guarantee the convergence. Conducted experiments on a real data set show its high performance when compared with the original k-means version.

Keywords : pattern recognition, global terrorism database, Manhattan distance, k-means clustering, terrorism data analysis
Conference Title : ICMLA 2017 : International Conference on Machine Learning and Applications
Conference Location : Copenhagen, Denmark
Conference Dates : June 11-12, 2017