HILARIS
Journal of Computer Science & Systems Biology
ISSN: 0974-7230
2020
Vol.13 No.6

Two New Algorithms, Critical Distance Clustering and Gravity Center Clustering

Farag Hamed Kuwil

Karabuk University, Turkey

Abstract

We developed a new algorithm based on Euclidean distance among data points and employing some mathematical statistics operations and called it critical distance clustering (CDC) algorithm (Kuwil, Shaar, Ercan Topcu, & Murtagh, Expert Syst. Appl., 129 (2019) 296–310. https://authors.elsevier.com/a/1YwCc3PiGTBULo). CDC works without the need of specifying parameters a priori, handles outliers properly and provides thorough indicators for clustering validation. Improving on CDC, we are on the verge of building second generation algorithms that are able to handle larger size objects and dimensions dataset.

Our new unpublished Gravity Center Clustering (GCC) algorithm falls under partition clustering and is based on gravity center "GC" and it is a point within cluster and verifies both the connectivity and coherence in determining the affiliation of each point in the dataset and therefore, it can deal with any shape of data, lambda is used to determine the threshold and identify the required similarity inside clusters using Euclidean Distance. Moreover, two coefficients lambda and n provide to the observer some flexibility to control over the results dynamically (parameters and coefficients are different, so, in this study, we assume that existing parameters to implement an algorithm as disadvantage or challenge, but existing coefficient to get better results as advantage), where n represents the minimum number of points in each cluster and lambda is utilized to increase or decrease number of clusters. Thus, lambda and n are changed from the default value in case of addressing some challenges such as outliers or overlapping.

Biography:

Farag Kuwil is a PhD student in Karabuk University. He has invented two data clustering algorithms. One published in Expert System with Application journal and the second in the first round of review

Speaker Publications:

1. A. K. Jain, "Data clustering: 50 years beyond K-means," Pattern Recognition Letters, vol. 31, no. 8, pp. 651–666, 2010.
View at: Publisher Site | Google Scholar
2. Z. Xie, R. Dong, Z. Deng et al., "A probabilistic approach to latent cluster analysis," in Proceedings of the Twenty Third International Joint Conferences on Artificial Intelligence (IJCAI), Beijing, China, August 2013.
View at: Google Scholar
3. T. VoVan and T. Nguyen Trang, "Similar coefficient of cluster for discrete elements," Sankhya B, vol. 80, no. 1, pp. 19–36, 2018.
View at: Publisher Site | Google Scholar
4. J. MacQueen, "Some methods for classification and analysis of multivariate observations," in Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, pp. 281–297, Oakland, CA, USA, June 1967.

7th International Conference on Big Data Analysis and Data Mining - July 17-18, 2020 Webinar.

Abstract Citation:

Farag Kuwil, Two New Algorithms, Critical Distance Clustering and Gravity Center Clustering, Data Mining 2020, 7th International Conference on Big Data Analysis and Data Mining – July 17-18, 2020 Webinar.

(https://datamining.expertconferences.org/speaker/2020/farag-hamed-kuwil-karabuk-university-turkey)