

Holistic Non-Unique Clustering

ISSN 1751-3030

Author:

Ramesh Chandra Bagadi

Data Scientist

INSOFE (International School Of Engineering),

Hyderabad, India.

rameshcbagadi@uwalumni.com

+91 9440032711

Technical Note

Theory

Given M number of points, each belonging to R^N , we first find the Proximity Matrix P_{ij} for each (M number of) point with each of all (M Number of points) points, inclusive of itself. The Proximity can be found using Euclidean distance or using the concept stated in [1]. We now find the Proximity Contrast Ratio $\delta = \frac{\text{Min}(P_{ij})}{\text{Max}(P_{ij})}$.

Now, for each

P_{ij} , we find all points that are within

$P_{ij} + \delta(P_{ij})$

distance for each point $\bar{x} \in R^N$, i.e., we find all the points that have connectivity among themselves (at least to one point among themselves) within the distance $P_{ij} + \delta(P_{ij})$. All these points can be called to belong to a given cluster. In this fashion, we can find M^2 number of overlapping Clusters where the membership of a point may not be unique to a given Cluster. We call this type of Clustering as Holistic Non-Unique Clustering.

References

1. Bagadi, R. (2017). Using the Appropriate Norm In The K-Nearest Neighbours Analysis. ISSN 1751-3030. *PHILICA.COM Observation number 173*.
http://www.philica.com/display_observation.php?observation_id=173
2. <http://www.philica.com/advancedsearch.php?author=12897>
3. http://www.vixra.org/author/ramesh_chandra_bagadi