Holistic Non-Unique Clustering
ISSN 1751-3030

Author:
Ramesh Chandra Bagadi
Data Scientist
INSOFE (International School Of Engineering),
Hyderabad, India.
rameshcbagadi@uwalumni.com
+91 944032711

Technical Note

Abstract

In this research technical Note the author have presented a novel method to find all Possible Clusters given a set of points in N Space.

Theory

Given $M$ number of points $\bar{x}_i \in \mathbb{R}^N$, $i = 1$ to $M$, each belonging to $\mathbb{R}^N$, 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 $\delta_{\text{Min}}^{\text{Max}} = \frac{\text{Min}(P_{ij})}{\text{Max}(P_{ij})}$ with only those values of $P_{ij} \neq 0$. Now, we consider any $P(i,j)$ which are $\left(\frac{M^2 - M}{2}\right)$ in number as The Proximity Matrix is Symmetric and also all the diagonal elements are equal to zero, and compute the distance $d\left\{P(i,j), \delta_{\text{Min}}^{\text{Max}}\right\} = P(i,j) + \left(\delta_{\text{Min}}^{\text{Max}}\right)(P(i,j))$. Now, we consider any point $\bar{x}_i \in \mathbb{R}^N$ and find all points (inclusive of $\bar{x}_i$) that have at least one neighbouring point within the distance $d\left\{P(i,j), \delta_{\text{Min}}^{\text{Max}}\right\}$, considered among themselves. We say that all such points form one Cluster. In this fashion, we can find at most $\left(\frac{M^2 - M}{2}\right)$ 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. Also, we can consider, all possible Proximity Contrast Ratio’s among the \( \frac{M^2 - M}{2} \) number of unique elements in the Proximity Matrix and can get at most \( \frac{M^2 - M}{2} \) number of overlapping Clusters for each of the \( \binom{M^2 - M}{2} \) \( C_2 \) number of possible Proximity Contrast Ratio’s Possible. Therefore, we can see at most \( \binom{M^2 - M}{2} \) \( C_2 \) \( \frac{M^2 - M}{2} \) number of clusters for the given Set of M Points.

**References**