Multi Class Classification Using 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

Abstract

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

Theory

Definition of a Cluster

We define a Cluster as follows:

A Cluster is a collection of Points (or objects) wherein they are scattered (their property is distributed) in such a fashion that, for a specified distance (measured in appropriate Metric of concern using appropriate Norm of concern) every point of this cluster has at least one neighbouring point also belonging to this cluster located within

(i) this specified distance* [1]
(ii) a certain small neighbourhood of this this specified distance, measured from the aforementioned point of concern.

Proximity Ratio

Given M number of points $\bar{x}_i \in R^N$, $i = 1 \; to \; M$, each belonging to R^N, we find the Proximity Matrix P for each (M number of) point with each of all other (M Number of points) points, inclusive of itself. The Proximity can be found using Euclidean distance or using the concept stated in [1].

$$P = \begin{bmatrix} d(1,1) & d(1,2) & d(1,3) & .\;.\;.\;. & d(1,(m-1)) & d(1,m) \\ d(2,1) & d(2,2) & d(2,3) & .\;.\;.\;. & d(2,(m-1)) & d(2,m) \\ d(3,1) & d(3,2) & d(3,3) & .\;.\;.\;. & d(3,(m-1)) & d(3,1) \\ . & . & . & .\;.\;.\;. & . & . \\ . & . & . & .\;.\;.\;. & . & . \\ . & . & . & .\;.\;.\;. & . & . \\ . & . & . & .\;.\;.\;. & . & . \\ d((m-1),1) & d((m-1),2) & d((m-1),3) & .\;.\;.\;. & d((m-1),(m-1)) & d((m-1),m) \\ d(m,1) & d(m,2) & d(m,3) & .\;.\;.\;. & d(m,(m-1)) & d(m,m) \end{bmatrix}$$

Proximity Contrast Ratio

We define the Proximity Contrast Ratio as $\delta_{\frac{P(p,q)}{P(l,m)}} = \frac{P(p,q)}{P(l,m)}$ with only those values of $P(p,q) \neq 0$ and $P(l,m) \neq 0$ with $p,q,l,m = 1\, to M$. Furthermore, $p \neq l$ and $q \neq m$ simultaneously.

Proximity Full Contrast Ratio

We now define the *Proximity Full Contrast Ratio* $\delta_{\frac{Min}{Max}} = \frac{Min(P(i,j))}{Max(P(i,j))}$ with only those values of $P(i,j) \neq 0$. Also, $i,j = 1\, to M$.

Cluster Analysis Based On Proximity Full Contrast Ratio

Case 1:

Now, we now compute the distance $d = +r\left(\delta_{\frac{Min}{Max}}\right)$ with $r = 1,2,3,4,5,...$ wherein $Max(r)$ gotten such that $Max(P_{ij}) = Max(r)\left(\delta_{\frac{Min}{Max}}\right)$.

Now, we consider any point $\bar{x}_i \in R^N$ and find all points that have at least one neighbouring point within the distance $d = +r\left(\delta_{\frac{Min}{Max}}\right)$, considered among themselves. That is, we find all the points which satisfy the Definition of a Cluster, afore-stated. We say that all such points form one Cluster each for each value of $r = 1,2,3,4,5,...$.

Case 2:

Now, we now compute the distance $d = -r\left(\delta_{\frac{Min}{Max}}\right)$ with $r = 1,2,3,4,5,...$ wherein $Max(r)$ gotten such that $Max(P_{ij}) = Max(r)\left(\delta_{\frac{Min}{Max}}\right)$.

Now, we consider any point $\bar{x}_i \in R^N$ and find all points that have at least one neighbouring point within the distance $d = -r\left(\delta_{\frac{Min}{Max}}\right)$, considered among themselves. That is, we find all the points which satisfy the Definition of a Cluster, afore-stated. We say that all such points form one Cluster each for each value of $r = 1,2,3,4,5,...$.

Cluster Analysis Based On Any Proximity Contrast Ratio

Case 1:

Now, we now compute the distance $d = +r\left(\delta_{\frac{P(p,q)}{P(l,m)}}\right)$ with $r = 1,2,3,4,5,...$ wherein $Max(r)$ gotten such that $Max(P_{ij}) = Max(r)\left(\delta_{\frac{P(p,q)}{P(l,m)}}\right)$.

Now, we consider any point $\bar{x}_i \in R^N$ and find all points that have at least one neighbouring point within the distance $d = +r\left(\delta_{\frac{P(p,q)}{P(l,m)}}\right)$, considered among themselves. That is, we find all the points which satisfy the Definition of a Cluster, afore-stated. We say that all such points form one Cluster each for each value of $r = 1,2,3,4,5,...$.

Case 2:

Now, we now compute the distance $d = -r\left(\delta_{\frac{P(p,q)}{P(l,m)}}\right)$ with $r = 1,2,3,4,5,...$ wherein $Max(r)$ gotten such that $Max(P_{ij}) = Max(r)\left(\delta_{\frac{P(p,q)}{P(l,m)}}\right)$.

Now, we consider any point $\bar{x}_i \in R^N$ and find all points that have at least one neighbouring point within the distance $d = -r\left(\delta_{\frac{P(p,q)}{P(l,m)}}\right)$, considered among themselves. That is, we find all the points which satisfy the Definition of a Cluster, afore-stated. We say that all such points form one Cluster each for each value of $r = 1,2,3,4,5,...$.

Advantages

This theory can be used for multi class classification, wherein an entity of concern may belong to more than one class. If we can also slate Belonging to a Cluster or Class ness based on grading of the Similarity between such affiliations, we can get a Model of Unique Holistic Clustering which can be used for Image Segmentation.

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