FUZZY CLUSTER

Cluster analysis divides data into groups (clusters) such that similar data objects belong to the same cluster and dissimilar data objects to different clusters. The resulting data partition improves data understanding and reveals its internal structure. Partitioning clustering algorithms divide up a data set into clusters or classes, where similar data objects are assigned to the same cluster whereas dissimilar data objects should belong to different clusters. In real applications there is very often no sharp boundary between clusters so that fuzzy clustering is often better suited for the data. Membership degree between zero and one are used in fuzzy clustering instead of crisp assignments of the data to clusters. The most prominent fuzzy clustering algorithm is the fuzzy c-means, a fuzzification of k-means.

Areas of application of fuzzy cluster analysis include for example data analysis, pattern recognition, and image segmentation. The detection of special geometrical shapes like circles and ellipses can be achieved by so-called shell clustering algorithms. Fuzzy clustering belongs to the group of soft computing techniques (which include neural nets, fuzzy systems, and genetic algorithms).

FUZZY CLUSTERING ALGORITHMS

Fuzzy clustering extends this notion to associate each partition with every cluster using a membership function. The output of such algorithms is a clustering, but not a partition. Given a set data } X x{ , x , , x = 1 2 … n , where k x , in general, is a vector p k 1k k2 kp x x[, x , … x,]∈ R

For all Nn k ∈ the problem of fuzzy clustering is to find a fuzzy pseudo partition and the associated cluster centers by which the structure of the data is represented as best as possible.

Some of the objective function-based clustering algorithms

- Fuzzy c-means algorithm (FCM): It generates spherical clusters of approximately the same size.
- Gustafson-Kessel algorithm (GK): It provides ellipsoidal clusters with approximately the same size. There are also axis-parallel variants of this algorithm. The algorithm can also be used to detect lines (to some extent).

- Gath-Geva algorithm (GG)/Gaussian mixture decomposition (GMD): This is used to generate ellipsoidal clusters with varying size. This algorithm has also axis-parallel variants. The algorithm can also be used to detect lines (to some Fuzzy Clustering extent).
- Fuzzy c-varieties algorithm (FCV): It is generally used for the detection of linear manifolds (infinite lines in 2D).
- Adaptive fuzzy c-varieties algorithm (AFC): It can be used to detect line segments in 2D data.
- Fuzzy c-shells algorithm (FCS): It is used for detection of circles (no closed form solution for prototypes).
  - Fuzzy c-spherical shells algorithm (FCSS): It detects circular patterns in data.
  - Fuzzy c-rings algorithm (FCR): It is also used for detection of circles.
  - Fuzzy c-quadric shells algorithm (FCQS): It is used for detection of rectangles (and variants thereof)