Fuzzy Clustering Algorithms – Review of the Applications

Jiamin Li; Harold W. Lewis
Department of Systems Science and Industrial Engineering
The State University of New York at Binghamton
Binghamton, NY, USA
jli272@binghamton.edu; hlewis@binghamton.edu

***Abstract*—Fuzzy clustering is an alternative method to conventional or hard clustering algorithms, which makes partitions of data containing similar subjects. The tendency of adopting machine learning, big data science, cloud computation in various industries depends on unsupervised learning on data structures to tell the story about consumers' behavior, fraud detection, and market segmentation. Fuzzy clustering contrasts with hard clustering by its nonlinear nature and discipline of flexibility in grouping massive data. It provides more accurate and close-to-nature solutions for partitions and herein implies more possibility of solutions for decision-making. In the specific matter of computation, fuzzy clustering has its roots in fuzzy logic and indicates the likelihood or degrees of one data point belonging to more than one group. This paper focuses on the study of models of fuzzy clustering in various cases. Uniquely designed algorithms enhance the accuracy of outcomes and are worth studying to assist future work. In some case scenarios, modeling processes are data-driven and place emphasis on the distances between points and new centers of clusters.In some other cases, which aim at market segmentation or evaluation of patients by healthcare records, membership degree is a key element in the algorithm. This paper surveys a wide-range of research that has well-designed mathematic models for fuzzy clustering, some of which include genetic algorithms and neural networks. The last section introduces open sources of Python and displays sample results from hands-on practice with these packages.**

Keywords—fuzzy c-mean clustering; pattern recognition; genetic algorithm; neural network; validity index

I. INTRODUCTION

Fuzzy clustering is a standalone type of unsupervised learning for classifying the patterns of datasets by investigating data structures. Companies and organizations capture data from cloud databases, machine-generated sensing data, and social media. These data, being captured or generated rapidly, are often referred to as big data, which could be structured, semi-structured or in a random format. These companies and organizations rely on data mining techniques to interpret these data, and clustering is one important section of data mining. There are two main categories of clustering, namely: hard and fuzzy type. The hard clustering groups data with distinguished boundaries, and forces each data point to belong to a specific group with the same pattern, which, in some cases, distorts the true value of data as well as limits the solutions to possible outcomes. Fuzzy clustering, on the other hand, proposes more diverse results.

II. FUZZY CLUSTERING ALGORITHMS

A. Basic Notions

- Data: Data can be numerical, categorical or mixed. Data in matrix form contains features and subjects in different units, such as time and value.
- Clusters: Cluster means a group of dataset or data points which share similarities. A mathematic interpretation of similarity is distance or distance norm. Data structure is the key for model clustering algorithms.
- Degree of membership: The degree of likelihood of one dataset belonging to several centers. The sum of membership degrees is equal to 1.

B. Fuzzy Partition, Data Structure, and Distance

Fuzzy clustering is a sophisticated technique for handling data which are unlabeled, contains outliers, and includes unusual patterns. Membership functions of fuzzy methods provide the possibility of one data point belonging to many groups, in some marketing applications, the study of overlap among groups is the core to explore business initiatives [1]. Fuzzy c-mean is one of the most widely applied and modified techniques in applications [2]. Data are generated by a possibility distribution or collected from various resources; Euclidean distance is the measurement used in most clustering algorithms to determine new centers [3]. In other cases, researchers design more precise distance equations, set up special variables or apply optimization accordingly under the content of cases and available data structure. Because Euclidean distance leads to clustering outcomes of spherical shapes, which is suitable for most cases, it is a top choice for many applications. Nevertheless, Gustafson-Kessel employed the Mahalanobis distance to determine different shapes of clusters, such as ellipsoidal cluster [4], and later on, Gath and Geva added maximum likelihood estimation to determine an inducing possibility of ellipsoidal shape and its size [5]. Motivated by rendering a more reasonable shape for outcomes, researchers developed fuzzy c varieties, adaptive fuzzy clustering, fuzzy c-mean, Gustafson-Kessel algorithm, and Gath-Geva algorithm [6].

In general applications, fuzzy clustering algorithms have been proved to be a better method than hard clustering in dealing with discrimination of similar structures [7], dataset in n-dimensional spaces [8], and is more useful for unlabeled data

978-1-5090-5263-9/16 $31.00 © 2016 IEEE
DOI 10.1109/SmartCloud.2016.14

and dataset with outliers [9]. Fuzzy c-mean has 66% accuracy in general cases, and Gustafson-Kessel scored 70% [10]. Fuzzy c-means proved to provide better solutions in machine learning, and image processing than hard clustering such as Ward's clustering and the k mean algorithm [11-16].

The weakness of fuzzy c-means is its sensitivity of outcome to the prototypes and also the optimizing process [17-19].

A fuzzy c-mean is a minimization function:

$$J_m = \sum_{i=1}^{N} \sum_{j=1}^{C} u_{ij}^m \left\| x_i - c_j \right\|^2, 1 \leq m \leq \infty \qquad (1)$$

where x_i is the i^{th} dataset or point in database, c_j is the j^{th} center assigned for cluster, $\|*\|$ means the distance between dataset or point to center. $\{m | m \in R > 1\}$ is a fuzziness index indicating ambiguity of an event and has its roots in the random concept of fuzzy logic. The objective function above needs the result from iterative algorithms below:

$$u_{ij} = \frac{1}{\sum_{k=1}^{C} \left(\frac{\left\| x_i - c_j \right\|}{\left\| x_i - c_k \right\|} ^{\frac{2}{m-1}} \right)} \qquad (2)$$

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^m . x_j}{\sum_{i=1}^{N} u_{ij}^m} \qquad (3)$$

u_{ij} is the degree of membership of individual x_i belonging to the cluster j; c_j is the center. Both of them are taken intuitively case by case.

The computation follows these steps:

- Initialization of the center of clusters. Compute the distance of each data point to the centers. (1) and (2) employs Euclidean distance.
- Compute membership degree of each point to each center. The sum of membership degree of all points equals 1.
- Based on the calculated membership degree compute the new centers of each cluster based on (3).
- Iterate these steps until the difference between J_m, c_j, or u_{ij} from the previous generation and current generation less than ε. ε is referred as the controller of fuzziness, which differs in cases, it could be an intuitive small number or a prior determined number [20-22].

In real-world applications, data are preferred to be presented by histograms, because they store information by empirical distributions which cover up some lost information and are more secure for users like banks [23]. Research about how to cluster histogram data set used L^2 Wasserstein distance [23-24] because this distance measurement is more precise for distribution property of histogram or other similar type of dataset.

A simple definition of Wasserstein distance is a measurement of distance between two possibility measures. Knowing each histogram has an associated quantile function, Wasserstein distance equals an integral power of variance between coupled quantile functions of dataset over an interval of 0 and 1. A simple definition of L^2 Wasserstein distance in the cluster algorithm is an equation combining Euclidean distance between means of two histograms with the Wasserstein distance between their centered quantile function [24].

The following shows a non-adoptive L^2 Wasserstein distance equation in the application of fuzzy c-mean for univariate histogram dataset:

$$d(y_i, g_k) = \sum_{j=1}^{p} \left(\bar{y}_{ij} - \bar{y}_{gj} \right)^2 + \sum_{j=1}^{p} d_W^2 \left(y_{il}^c, g_{kj}^c \right) \qquad (4)$$

This equation based on prototype g_k, which is a predetermined set of centered quantile functions, as well as a set of empirical distribution. The minimization subjective is defined as:

$$J(G, \mathrm{U}) = \sum_{k=1}^{K} \sum_{i=1}^{n} (u_{ik})^m \, d(y_i, g_k) \qquad (5)$$

Accordingly, the membership degree function is as follows:

$$u_{ik} = \left[\sum_{h=1}^{K} \left(\frac{d(y_i, g_k)}{d(y_i, g_h)} \right)^{\frac{1}{m-1}} \right]^{-1} \qquad (6)$$

The computation steps are the same with the basic fuzzy c-mean algorithm that have been stated earlier. A more adaptive method extends (4) by timing a matrix of vector weights in both parts of equation to indicate various importance of data point and adding a constraint to iterative algorithm (6) while getting the desired result.

Interval-valued data is another interesting data structure used in different places, such as banking, environment, and food industries [25-28]. An interval-valued data graphically can have two vertices which imply a lower bound and an upper bound of a variable of an object which has been observed at a time point. Mathematically, interval-valued data can be stored as a vector of two points, namely: midpoint of upper bound and lower bound, and a radius of how this number could float. Outliers could exist in either or both dimension. Matrix x_i contains midpoint and radius and its cardinality matrix $\tilde{x}_i$ have been designed for handling such outliers in some fuzzy clustering algorithms.

A minimization function is written as following:

$$min: \sum_{i=1}^{I} \sum_{c=1}^{C} u_{ic}^m \, {}_{exp}D^2(x_i, \tilde{x}_c) \approx \sum_{i=1}^{I} \sum_{c=1}^{C} u_{ic}^m \, [1 -$$

$$\exp\{-\beta(\|m_i - \widetilde{m_c}\|^2 + \|r_i - \widetilde{r_c}\|^2 \qquad (7)$$

D'Urso et al. proposed and used the squared exponential distance between two matrices x_i and $\tilde{x}_c$ [29-30]. The exponential distance has the nature of weights data points. It assigns small and larger weights to outliers and data points compact to others. Accordingly, the membership degree is computed by:

$$u_{ic} = \frac{1}{\sum_{c'=1}^{C}\left[\frac{exp^{D^2(x_i,\tilde{x}_c)}}{exp^{D^2(x_i,\tilde{x}_{c'})}}\right]^{\frac{1}{m-1}}} \quad (8)$$

The overall steps to solve the minimization function are the same as the ones for basic fuzzy c-mean, yet β is a crucial value to the entire algorithm which leads to more work on robust design and validation on such values [31-32].

Mahalanobis distance has been used most in applications of image processing. It measures a data point to a distribution [33-36]. It has a general form as following:

$$d_{ij} = \left(x_i - \mu_j\right)^T M\left(x_i - \mu_j\right) \quad (9)$$

M is equal to the inverse of the matrix of the j^{th} cluster. Fuzzy clustering models, such as Miyamato and Mukaidono, often use Mahalanobis distance in one or more parts of the algorithm. It is proved by research that Mahalanobis distance is a superior equation for 2D datasets [37-39].

C. *Validity Index*

The validity Index is an analytical tool for evaluating the performance of clustering algorithms. Validity indexes for hard clustering methods evaluate the boundaries, while such indexes evaluate the membership degree for the fuzzy clustering methods [40]. Validity indexes have the formula:

$$\max(\min) z = f(\Delta_c, \delta_c), c = 1,2,\ldots,C \quad (10)$$

Δ_c stands for compactness within cluster and is referred to as *intradistance*; δ_c which stands for the separation of clusters and is referred to as *interdistance*. The validity index is designed for minimizing the compactness and maximizing the separation [41].

There are two most widely-used validity indexes for fuzzy clustering, namely: partition entropy (PE) and the Xie and Beni et al.' index (XB):

$$V_{PE} = -\frac{1}{n}\sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij} log_a u_{ij} \quad (11)$$

c , number of centers, is computed in range $[1, log_a]$ in PE and has optimal value by minimizing V_{PE}.

$$V_{XB} = \frac{J_m(U,V)}{Sep(V)} = \frac{\sum_{i=1}^{c} u_{ij}^m \left\|x_j - v_i\right\|^2}{n \min_{i \neq j}\left\|v_i - v_j\right\|^2} \quad (12)$$

$J_m(U,V)$ indicates Δ_c, and $Sep(V)$ indicates δ_c. Through minimizing V_{XB}, c can be optimized.

More recent developments of the validity index, such as dual center and gap statistic, extend such an index for more diverse data structures [42-43].

D. *Fuzzy Clustering and Neural Networks*

Machine learning is a way to automate the process, such as image interpreting, data entry, and factory monitoring. Studies on radial basis function networks state an input-output relation [44-45] which connect with fuzzy clustering to process data structure in the input and output space [46]:

$$J_{OFC} = \sum_{k=1}^{n}\sum_{i=1}^{C}(\mu_{ik})^m \left\|y_j - v_y\right\|^2 \quad (13)$$

J_{OFC} stands for the fuzzy c-mean clustering for the output data structure.

$$v_i = \sum_{k=1}^{n}(\mu_{ik})^m y_k / \sum_{k=1}^{n}(\mu_{ik})^m \quad (14)$$

$$\mu_{ik} = 1/\sum_{j=1}^{c}\left(\frac{\|y_k - v_i\|}{\|y_k - v_j\|}\right)^{\frac{2}{m-1}} \quad (15)$$

v_i indicates the center, while $\mu_{ik}$is the membership degree. v_i is employed to determine data in input or produce space, where the input data has a two-dimensional vector in form of $[x_k^T, y_k]$:

$$J_{IOFC} = \sum_{k=1}^{n}\sum_{i=1}^{C}(\mu_{ik})^m\left(\|x_k - v_i\|^2 + \gamma\left\|y_j - v_y\right\|^2\right) \quad (16)$$

$$\gamma = \gamma_0 e^{-\frac{t}{\tau}} \quad (17)$$

$$\mu_{ik} = 1/\sum_{j=1}^{c}\left(\frac{\|x_k - v_i\|^2 + \gamma\|y_k - v_i\|^2}{\left\|x_k - v_j\right\|^2 + \gamma\left\|y_k - v_j\right\|^2}\right)^{\frac{2}{m-1}} \quad (18)$$

γ is the scaling factor, τ is the constant time, and t indicates the iteration. This process defines the membership degree in more specific to the data structure. Optimization function is another way to obtain validity index [46].

E. *Genetic Algorithms.*

Missing or incomplete data happens in many real-world cases. Quantity and quality of the missing or incomplete data lead to the issue of learning outcomes. Genetic algorithms have been a complement to optimization or as an additional part of fuzzy clustering algorithms [47-49].

$$X = \begin{bmatrix} x_{11} & \ldots & x_{1l} \\ \vdots & \ldots & \vdots \\ x_{n1} & \ldots & x_{nl} \end{bmatrix} \quad (20)$$

X is a matrix data set about traffic volume which has been collected in n time intervals on l days of week. While the overall steps for the modeling cluster stays the same as fuzzy c-mean, an estimation of the missing values of X has formulized as following:

$$\widehat{x_{ij}} = \sum_{k=1}^{k} U(x_i, c_k) \cdot c_k \tag{21}$$

This estimation is for the calculation of the root mean square error. Furthermore, it is used for goodness-of-fit measurement in genetic algorithm:

$$error(U, c) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{l}(x_{ij} - \widehat{x_{ij}})^2} \tag{22}$$

$$f(U, c) = \frac{1}{error(U,c) + constant} \tag{23}$$

The implementation of (22) and (23) for the error estimation and fitting process of genetic algorithms is for the optimization of the membership degrees and the centers of clusters. Many studies apply genetic algorithms for missing data and emphasize membership functions for solutions [50-53].

III. APPLICATIONS REVIEW

This section provides a table from a comprehensive review of the applications from relevant recent researches.

TABLE I.

Applied Industry	Description	Name of the Fuzzy algorithm	centers/Validity Index	Applied Distance Equation	
Finance	Interrelation with companies	Fuzzy c-mean	Dunn's	Euclidean distance	[54]
	Define interrelationship between portfolios	Fuzzy c-means	Xie-Beni validation index	Pearson correlation	[55]
Energy	Predict downtime of high-speed milling (HSM)	Sequential fuzzy c-mean dynamic	No limit	The Gaussian functions	[56]
	Sort the signal to optimize transmission	Optimal Fuzzy C-Means clustering	Subject	Prior determined	[57]
	Improve working time in wireless sensor network	Decentralized Fuzzy Clustering Protocol	Set to 5	Euclidean distance	[58]
	Improve lifetime	Multi-subjective fuzzy clustering	Subject	Received signal strength	[59]
Medicine	GA for getting optimize parameters for fuzzy c-mean	Neighborhood intuitionistic fuzzy c-means clustering algorithm with a genetic algorithm	Gray matrix transformed from medical image	Euclidean distance	[60]
	Machine learning	Fuzzy c-means	Subject	Ahmad and Dey	[61]
	Optimize the cluster center for image processing	Image segmentation algorithm	Davies–Bouldin; Xie; Beta; Dunn	Euclidean distance	[62]
Web classification	Pattern discovering from web logins	Fuzzy means	Ratio of compactness	Euclidean distance	[63]
	Machine learning	Fuzzy c-means	Categorical terms	Euclidean distance	[64]
	Web crawler cluster	Potential-Based Clustering Algorithm	v = COMP/SEP	Hamming Distance	[65]
Health care	extract the knowledge from information	Mixed Fuzzy Clustering (MFC) algorithm	Subject	Euclidean distance	[66]
	Analyze text format medical data	Fuzzy c-means algorithm	None	Frequency matrix	[67]
	Context selection from Body Sensor Networks	Fuzzy c-means algorithm with GA optimization	GA elicit selection	Euclidean distance	[68]
	Determine the center of cluster with miss data	Fuzzy c-means optimal completion strategy (OCS)	Historical data	Minkowski distance	[69]
Marketing	Extract revenue and usage pattern from customer	Fuzzy c-means	Xie and Beni's	Euclidean distance	[70]
	electricity consumption and demography information	Mixed Fuzzy Clustering	Calinski-Haabasz, Davies-Bouldin; Silhouette index	Euclidean distance with λ weights for spatial data	[71]
Big Data	Applied real data stream to determine storage system	Fuzzy Incremental Clustering Approach	Subject	Euclidean distance	[72]
	New structure algorithm for large scale data	Density-based weighted FCM algorithm	Average adjusted Rand index (ARI)	Euclidean Distance	[73]
	Particle Swarm Optimization for big data	Fuzzy c-means algorithm	Particle Swarm Optimization	Euclidean distance	[74]
Machine Learning	Determine the membership function based on the data with noises and outliers	Partition index maximization (PIM) clustering	X.L. Xie, G. Beni	Subject	[75]
	From sensor data and updating training data detect events	Adjustable fuzzy clustering algorithm (AFC)	A set of training samples	Merging-mechanism	[76]
Pattern recognition	Optimize the network in support	Type-2 fuzzy clustering	Given numbers	None	[77]
	Pedestrian detection from infrared image	Adaptive fuzzy C-means clustering	Image intensity	None	[78]
Time-series prediction	Modeling based on time series data	Hybrid FCM and Fuzzy C Medoids technique	Fuzzy C Medoids	Dynamic Time Warping Distance	[79]

	Extract training patterns based on TSK fuzzy rule	Incremental clustering algorithm with TSK Fuzzy rule	None	Hybrid distances	[80]
	Use fuzzy clustering technique to determine membership value	Gustafson-Kessel fuzzy clustering	$2 \leq c \leq n$	Mahalanobis distance	[81]
Robust Design	Designed goal aim at the handle outliers and interval-valued data	Trimmed Fuzzy C-medoids for interval-valued data (TrFCMd-ID)	Fuzzy Rand index	Euclidean distance	[82]

IV. OPEN SOURCE FOR FUZZY C-MEAN

There are many open sources based on Python relating to fuzzy clustering, such as skfuzzy, sklearn.cluster which are available from Github as open source for learning and virtualize fuzzy partitions [83][84]. Samples shown in Fig.1 and Fig.2.

V. CONCLUSION

Although research has developed many clustering algorithms, various structures of data are causing problems for adapting those well-constructed models. The common approaches are based on the cooperation of sophisticated distance equations, utilization of validity indexes, collaboration with other algorithms to obtain more accurate membership degrees or centers. The table in the previous section summarizes strategies for different data structures provides insight on how to manipulate basic fuzzy c-mean and upgrade the algorithm. The motivation for future work could look into new optimization functions and hybrid algorithms to enhance the overall process.

REFERENCES

[1] Fung, G. (2001). A Comprehensive Overview of Basic Clustering Algorithms.

[2] Kaymak, U. and Setnes, M. (2000). Extended Fuzzy Clustering Algorithm. ERIM Report Series Research in Management. 1-23.

[3] Dunn, J. (1973). A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact, Well-Separated Cluster. Journal of Cybernetics. 3(3). 32-57.

[4] E.E. Gustafson, W.C. Kessel, Fuzzy Clustering with a Fuzzy Covariance Matrix, IEEE CDC, San Diego, California, 1979, pp. 761–766.

[5] Gath, I. and Geva, A.B. (1989). Unsupervised Optimal Fuzzy Clustering. IEEE Transactions on Pattern Analysis and Machine Intelligence. 11(7). 773-781.

[6] Martin, E. (2003). Pap-Smear Classification. Technical the University of Denmark.

[7] Bayley, M.J., Gillet, V.J., Willett, P., Bradshaw, J. and Green, D.V.S. (1999). Computational Analysis of Molecular Diversity for Drug Discovery. Proceeding of the 3rd Annual Conference on Research in Computational Molecular Biology. ACM Press. New York. 321-330.

[8] Barnard, J. M., and Downs, G.M. (1992). Clustering of Chemical Structures on the Basis of Two-Dimensional Similarity Measures. Journal of Chemical Information and Computer Science. 32. 644-649.

[9] Feher, M. and Schmidt, J.M. (2003). Fuzzy Clustering as a Means of Selecting Representative Conformers and Molecular Alignment". Journal of Chemical Information and Computer Science. 43. 810-818.

[10] Guthke, R., Schmidt-Heck, W., Hahn, D. and Pfaff, M. (2002). Gene Expression data Mining for Functional Genomics using Fuzzy Technology. Advances in Computational Intelligence and Learning Methods and Applications. Kluwer. 475-487.

[11] Rodgers S.L., Holliday J.D. and Willet P. (2004). Clustering Files of Chemical Structures Using the Fuzzy k-Means Clustering Method. Journal of Chemical Information and Computer Science. 44. 894-902.

[12] Rastogi, Rohit, et al. "GA-Based Clustering of Mixed Data Type of Attributes (Numeric, Categorical, Ordinal, Binary, and Ratio-Scaled)." *BVICAM's International Journal of Information Technology* 7.2 (2015).

[13] Rezaee, M. Ramze, B.p.f. Lelieveldt, and J.h.c. Reiber. "A New Cluster Validity Index for the Fuzzy C-mean." *Pattern Recognition Letters* 19.3-4 (1998): 237-46. Web.

[14] J. C. Dunn, "A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters," J. Cybern., vol. 3, no. 3, pp. 32–57, 1973

[15] L. J. Hubert and P. Arabie, "Comparing Partitions," J. Classification, vol. 2, pp. 193–218, 1985.

[16] Bezdek, J. C. (1975). Mathematical models for systematics and taxonomy. In G. F. Estabrook, editor, Proceedings of the 8th International Conference on Numerical Taxonomy, San Francisco. Freeman

[17] Schwämmle, Veit, and Ole N. Jensen. "A Simple and Fast Method to Determine, the Parameters for Fuzzy c, means Cluster Validation." *arXiv preprint arXiv:1004.1307* (2010).

[18] Bezdek, J.C. (1974). Cluster Validity with Fuzzy Sets. J. Cybernet., 58-73. Bezdek, J.C.

[19] L. Zadeh, Fuzzy sets, Inform. Control 8 (1965) 338–353.

[20] J.C. Bezdek, Pattern Recognition with Fuzzy Subjective Function Algorithms, Kluwer Academic Publishers, Norwell, MA, USA, 1981.

[21] D.L. Pham, J.L. Prince, An Adaptive Fuzzy c-means Algorithm for Image Segmentation in the Presence of Intensity Inhomogeneities, Pattern Recogn. Lett. 20 (1999) 57–68.

[22] J. C. Dunn (1973): "A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters," *Journal of Cybernetics* 3: 32-57

[23] Irpino, Antonio, Rosanna Verde, and Francisco de-AT De Carvalho. "Dynamic Clustering of Histogram Data Based on Adaptive Squared Wasserstein distances." *Expert Systems with Applications* 41.7 (2014): 3351-3366.

[24] De Carvalho, Francisco de-AT, Antonio Irpino, and Rosanna Verde. "Fuzzy Clustering of Distribution-valued Data Using an Adaptive L 2 Wasserstein distance." *Fuzzy Systems (FUZZ-IEEE), 2015 IEEE International Conference on*. IEEE, 2015.

[25] Denoeux, T., & Masson, M. (2000). Multidimensional Scaling of Interval-valued Dissimilarity Data. Pattern Recognition Letters, 21(1), 83–92

[26] Coppi, R., & D'Urso, P. (2002). Fuzzy k-means clustering models for triangular fuzzy time trajectories. Statistical Methods and Applications, 11(1), 21–40.

[27] Guru, D. S., Kiranagi, B. B., & Nagabhushan, P. (2004). Multivalued type proximity measure and concept of mutual similarity value useful for clustering symbolic patterns. Pattern Recognition Letters, 25(10), 1203–1213.

[28] D'Urso, P., De Giovanni, L., & Massari, R. (2015b). Trimmed fuzzy clustering for interval-valued data. Advances in Data Analysis and Classification, 9(1), 21–40.

[29] D'Urso, P., & Giordani, P. (2006). A robust fuzzy k-means clustering model for interval valued data. Computational Statistics, 21(2), 251–269.

[30] Wu, K. L., & Yang, M. S. (2002). Alternative c-means clustering algorithms. Pattern Recognition, 35(10), 2267–2278.

[31] Krishnapuram, R., Joshi, A., Nasraoui, O., & Yi, L. (2001). Low-complexity fuzzy relational clustering algorithms for web mining. IEEE Transactions on Fuzzy Systems, 9(4), 595–607.

[32] D'Urso, Pierpaolo, et al. "Exponential distance-based fuzzy clustering for interval-valued data." *Fuzzy Optimization and Decision Making* (2016): 1-20.

[33]] R. Krishnapuram, J. Kim, A note on the Gustafson–Kessel and adaptive fuzzy clustering algorithms, IEEE Trans. Fuzzy Syst. 7 (1999) 453–461.

[34] D.E. Gustafson, W.C. Kessel, Fuzzy clustering with a fuzzy covariance matrix, in IEEE Conference on Decision and Control including the 17th Symposium on Adaptive Processes, vol. 17, San Diego, CA, USA, 1978, pp. 761–766.

[35] I. Gath, A. Geva, Unsupervised optimal fuzzy clustering, IEEE Trans. Pattern Anal. Mach. Intell. 11 (1989) 773–780.

[36]] S.R. Kannan, R. Devi, S. Ramathilagam, K. Takezawa, Effective FCM noise clustering algorithms in medical images, Comput. Biol. Med. 43 (2) (2013) 73–83.

[37] Fukui, Ken-ichi, et al. "Evolutionary distance metric learning approach to semi-supervised clustering with neighbor relations." *Tools with Artificial Intelligence (ICTAI), 2013 IEEE 25th International Conference on*. IEEE, 2013.

[38] Chen, Songcan, and Daoqiang Zhang. "Robust image segmentation using FCM with spatial constraints based on new kernel-induced distance measure." *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on* 34.4 (2004): 1907-1916.

[39] Benaichouche, A. N., Hamouche Oulhadj, and Patrick Siarry. "Improved spatial fuzzy c-means clustering for image segmentation using PSO Initialization, Mahalanobis Fistance, and Post-segmentation Correction."*Digital Signal Processing* 23.5 (2013): 1390-1400.

[40] J. C. Bezdek, "Cluster validity with fuzzy sets," J. Cybern., vol. 3, no. 3, pp. 58–73, 1973.

[41] J. C. Dunn, "A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters," J. Cybern., vol. 3, no. 3, pp. 32–57, 1973.

[42] Wu, Chih-Hung, et al. "A new fuzzy clustering validity index with a median factor for centroid-based clustering." *Fuzzy Systems, IEEE Transactions on*23.3 (2015): 701-718.

[43] Yue, Shihong, et al. "A new validity index for evaluating the clustering results by partitional clustering algorithms." *Soft Computing* (2015): 1-12.

[44] S. Haykin, Neural Networks: A Comprehensive Foundation, Prentice Hall, NJ, 1999

[45] J. Moody, C.J. Darken, Fast learning in networks of locally-tuned processing units, Neural Comput. 1 (1989) 281–294

[46] Tsekouras, George E., and John Tsimikas. "On training RBF neural networks using input–output fuzzy clustering and particle swarm optimization." *Fuzzy Sets and Systems* 221 (2013): 65-89

[47] Boyles S., 2011. A comparison of interpolation methods for missing traffic volume data. In: Proceedings of the 90th Annual Meeting of the Transportation Research Board, January, Washington, DC, pp.23–27.

[48] Chen, J., Shao, J., 2000. Nearest neighbor imputation for survey data. J. Official Stat. 16 (2), 113–131.

[49] Zhong, M., Lingras, P., Sharma, S., 2004. Estimation of missing traffic counts using factor, genetic, neural, and regression techniques. Transport. Res. Part C 12, 139–166.

[50] Di Nuovo, A.G., 2011. Missing data analysis with fuzzy C-means: a study of its application in a psychological scenario. Expert Syst. Appl. 38 (6), 6793–6797

[51] Hathaway, R.J., Bezdek, J.C., 2002. Clustering incomplete relational data using the non-Euclidean relational fuzzy c-means algorithm. Pattern Recogn. Lett. 23, 151–160.

[52] Liao, Z., Lu, X., Yang, T., Wang, H., 2009. Missing data imputation: a fuzzy K-means clustering algorithm over sliding window. In: Proceedings of the 6th International Conference on Fuzzy Systems Knowledge Discovery, Tanjin, August, pp. 133–137.

[53] Tang, Jinjun, et al. "A hybrid approach to integrate fuzzy C-means based imputation method with genetic algorithm for missing traffic volume data estimation." *Transportation Research Part C: Emerging Technologies* 51 (2015): 29-40.

[54] Tufan E, Hamarat B: Clustering of financial ratios of the quoted companies through fuzzy logic method. J Naval Sci Engin, 1 (2): 123-140, 2003.

[55] A. Stetco, X. Zeng, J. Keane, Fuzzy cluster analysis of financial time series and their volatility assessment, in: Proceedings of 2013 IEEE International Conference on Systems, Man, and Cybernetics, 2013, pp. 91–96.

[56] Jahromi, Amin Torabi, et al. "Sequential fuzzy clustering based dynamic fuzzy neural network for fault diagnosis and prognosis." Neurocomputing(2016).

[57] Maity, Santi P., Subhankar Chatterjee, and Tamaghna Acharya. "On Optimal Fuzzy C-means Clustering for Energy Efficient Cooperative Spectrum Sensing in Cognitive Radio Networks." Digital Signal Processing 49 (2016): 104-15. Web.

[58] Alia, O.M. A Decentralized Fuzzy C-Means-Based Energy-Efficient Routing Protocol for Wireless Sensor Networks. Sci. World J. 2014.

[59] Sert, Seyyit Alper, Hakan Bagci, and Adnan Yazici. "MOFCA: Multi-subjective fuzzy clustering algorithm for wireless sensor networks." *Applied Soft Computing* 30 (2015): 151-165.

[60] Huang, Ching-Wen, et al. "Intuitionistic fuzzy $ $ c $ $ c-means clustering algorithm with neighborhood attraction in segmenting medical image." Soft Computing-A Fusion of Foundations, Methodologies, and Applications 19.2 (2015): 459-470.

[61] Ahmad, Amir. "Evaluation of Modified Categorical Data Fuzzy Clustering Algorithm on the Wisconsin Breast Cancer Dataset." Scientifica 2016 (2016).

[62] Bose, Ankita, and Kalyani Mali. "Fuzzy-Based Artificial Bee Colony Optimization for gray image segmentation." Signal, Image and Video Processing (2016): 1-8.

[63] Ansari, Zahid, Mohammad Fazle Azeem, A. Vinaya Babu, and Waseem Ahmed. "A Fuzzy Clustering Based Approach for Mining Usage Profiles from Web Log Data." (IJCSIS) International Journal of Computer Science and Information Security 9.6 (2011): 70-79. Web.

[64] Cosma, Georgina, and Giovanni Acampora. "A Computational Intelligence Approach to Efficiently Predicting Review Ratings in E-commerce." Applied Soft Computing (2016).

[65] Tsekouras, George E., and Damianos Gavalas. "An Effective Fuzzy Clustering Algorithm for Web Document Classification: A Case Study in Cultural Content Mining." Int. J. Soft. Eng. Know. Eng. International Journal of Software Engineering and Knowledge Engineering 23.06 (2013): 869-86. Web.

[66] Ferreira, Marta C., et al. "Fuzzy modeling based on Mixed Fuzzy Clustering for healthcare applications." Fuzzy Systems (FUZZ-IEEE), 2015 IEEE International Conference on. IEEE, 2015.

[67] A. Karami, A. Gangopadhyay, B. Zhou and H. Karrazi, "FLATM: A fuzzy logic approach topic model for medical documents," Fuzzy Information Processing Society (NAFIPS) held jointly with 2015 5th World Conference on Soft Computing (WConSC), 2015 Annual Conference of the North American, Redmond, WA, 2015, pp. 1-6.

[68] Fenza, Giuseppe, Domenico Furno, and Vincenzo Loia. "Hybrid approach for context-aware service discovery in healthcare domain." Journal of Computer and System Sciences 78.4 (2012): 1232-1247.

[69] Ben-Arieh, David, and Deep Kumar Gullipalli. "Data Envelopment Analysis of clinics with sparse data: Fuzzy clustering approach." Computers & Industrial Engineering 63.1 (2012): 13-21.

[70] Bose, Indranil, and Xi Chen. "Detecting the migration of mobile service customers using fuzzy clustering." Information & Management 52.2 (2015): 227-238.

[71] Schafer, Hanna, et al. "Analyzing the segmentation of energy consumers using mixed fuzzy clustering." Fuzzy Systems (FUZZ-IEEE), 2015 IEEE International Conference on. IEEE, 2015.

[72] Gaceanu, Radu D., and Horia F. Pop. "A fuzzy incremental clustering approach to hybrid data discovery." Acta Electrotechnica et Informatica 12.2 (2012): 16.

[73] Li, Yangyang, et al. "A study of large-scale data clustering based on fuzzy clustering." Soft Computing (2015): 1-12.

[74] Xianfeng, Yang, and Liu Pengfei. "Tailoring Fuzzy C-Means Clustering Algorithm for Big Data Using Random Sampling and Particle Swarm Optimization." IJDTA International Journal of Database Theory and Application 8.3 (2015): 191-202. Web.

[75] Wu, Zhenning, Huaguang Zhang, and Jinhai Liu. "A fuzzy support vector machine algorithm for classification based on a novel PIM fuzzy clustering method." Neurocomputing 125 (2014): 119-124.

[76] Wang, Zhelong, Ming Jiang, Yaohua Hu, and Hongyi Li. "An Incremental Learning Method Based on Probabilistic Neural Networks and Adjustable Fuzzy Clustering for Human Activity Recognition by Using Wearable Sensors." IEEE Transactions on Information Technology in Biomedicine IEEE Trans. Inform. Technol. Biomed. 16.4 (2012): 691-99. Web.

[77] Khormali, Aminollah, and Jalil Addeh. "A novel approach for recognition of control chart patterns: Type-2 fuzzy clustering optimized support vector machine." ISA Transactions (2016).

[78] John, Vijay, et al. "Pedestrian detection in thermal images using adaptive fuzzy C-means clustering and convolutional neural networks." Machine Vision Applications (MVA), 2015 14th IAPR International Conference on. IEEE, 2015.

[79] Izakian, Hesam, Witold Pedrycz, and Iqbal Jamal. "Fuzzy clustering of time series data using dynamic time warping distance." Engineering Applications of Artificial Intelligence 39 (2015): 235-244.

[80] Peng, Hung-Wen, et al. "Time series forecasting with a neuro-fuzzy modeling scheme." Applied Soft Computing 32 (2015): 481-493.

[81] Cagcag Yolcu, Ozge. "A hybrid fuzzy time series approach based on fuzzy clustering and artificial neural network with single multiplicative neuron model." Mathematical Problems in Engineering 2013 (2013)

[82] D'Urso, Pierpaolo, Livia De Giovanni, and Riccardo Massari. "Trimmed fuzzy clustering for interval-valued data." Advances in Data Analysis and Classification 9.1 (2015): 21-40.

[83] "Web Scale K-Means clustering" D. Sculley, *Proceedings of the 19th international conference on World wide web* (2010)

[84] Warner, Josh. Sciki-fuzzy. Program documentation. Github.Python, 2013.Web.

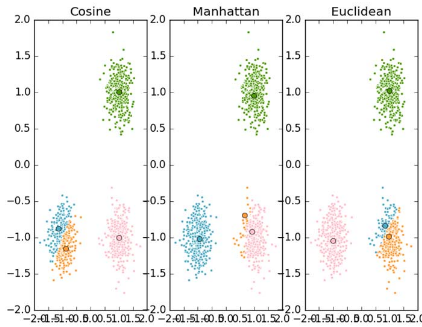

Figure 1: The application of fuzzy c-mean with 1000 random generated points and different distance equations.

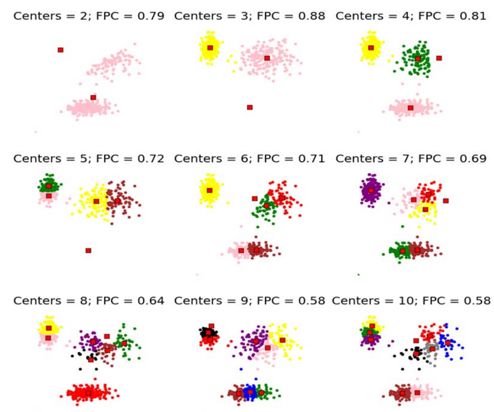

Figure 2: An examination of the centroid within 1000 randomly generated points for fuzzy c-mean.