Adaptive Reliability-Based Fuzzy Clustering with Modified Objective Function

Erind Bedalli[1*], Shkelqim Hajrulla[2]

[1]Department of Informatics, University of Elbasan 'Aleksander Xhuvani', Elbasan, Albania
[2]Department of Computer Engineering, Epoka University, Tirana, Albania

Abstract

Clustering is an unsupervised learning form which aims at revealing intrinsic patterns in data based on the distances or similarities of instances among each other. In real world datasets, clustering procedures may be significantly affected by noise, outliers and low-confidence data points. This paper presents a fuzzy clustering approach based on modified objective functions employing adaptive reliability measures, striving to enhance the robustness of the clustering results. The modification on the objective function integrates the notion of a point's influence into the clustering procedure. Therefore, the influence of each point will be controlled by a reliability score which will be evaluated adaptively. The root mean square propagation (RMSprop) algorithm is applied to dynamically assess the reliability score of the data points. Finally, this framework will be experimentally tested on several benchmark datasets to assess the quality of generated clusters and compare it to classical fuzzy clustering algorithms.

Keywords
fuzzy clustering, reliability score, objective function modifications, RMSprop

1. Introduction

In machine learning, clustering is a core unsupervised learning task which intends to group similar data instances together, forming clusters (subsets) where elements share strong similarities with one another and clear dissimilarities from those in other clusters. The clustering process is guided merely by the inter-point similarity or distance, without any available external information about the latent structure of the dataset. Clustering has proven to be a valuable and flexible technique due to its wide-ranging applications, including recommendation systems in e-commerce, customer profiling in marketing, topic discovery in text mining, behavioral pattern analysis in psychology, and species categorization in biological sciences [1]. For any sizes of datasets, clustering constitutes a powerful tool for exploring and summarizing data, but especially for larger volumes of data it remains essential for uncovering patterns, revealing hidden structures, and often guiding downstream machine learning tasks [2].

There are multiple approaches to the clustering problem, among which can be distinguished hard clustering and fuzzy clustering. In the hard clustering approach, each data instance is assigned to exclusively one cluster, consequently the cluster boundaries are crisp and without overlapping. In contrast, fuzzy clustering allows instances to belong to multiple clusters simultaneously with varying degrees of membership (values between 0 and 1), making it a more flexible and realistic approach in circumstances where data points are not clearly separable. The traditional fuzzy clustering algorithm operates by implying an equal contribution of the data points to the clustering procedure, nonetheless this makes the algorithm susceptible to presence of noise and outliers [3, 4].

[1]Proceedings of RTA-CSIT 2025, May 22-24, 2025, Tirana, Albania
*Corresponding author.
erind.bedalli@uniel.edu.al(E. Bedalli); shhajrulla@epoka.edu.al (S. Hajrulla)
0000-0002-1107-0165(E. Bedalli); 0000-0003-0803-8268 (S. Hajrulla)

This paper presents a modification to the classical fuzzy clustering algorithm, incorporating adaptive reliability scores in the definition of the objective function of the algorithm. The central idea of this modification is to regulate the influence of the data points into the clustering process, based on a dynamically assessed reliability score. More specifically, the proposed method computes the reliability score of each point using entropy-based confidence metrics and the entire model is trained using the RMSprop optimizer, in order to adaptively scale learning rates. This approach allows the algorithm to down-weight the influence of uncertain or noisy data points during the clustering procedure, avoiding distortions that they may induce to the generated clusters. As a result, the model strives to be more resilient towards the presence of outliers and operates more effectively upon real-world, imperfect datasets. Together, the use of reliability scoring and adaptive optimization are expected to contribute to more stable, accurate, and interpretable clustering results. This approach will be tested extensively on several benchmark datasets and slightly distorted variants of benchmark datasets to assess quantitively the quality of the generated clusters.

2. Related work

In the recent decades, significant progress has been made in the field of unsupervised learning, designing various forms of clustering approaches and enhancements tailored to specific scenarios. Nevertheless, clustering remains a challenging problem, there is no algorithm capable of learning the patterns of every dataset. The idea of modifying the objective function and the idea of reliability measures usage have been presented in different flavors in various research works, often separately, but also occasionally intertwined together. In this section, the main approaches in these directions will be summarized and the differences in our approach will be accentuated.

F. Hoppner et al. have explored how the traditional fuzzy c-means objective function can be modified to support different levels of fuzziness in cluster memberships. They have analyzed how these modifications of the objective function allow for more flexible, partial membership values, enabling the clustering algorithm to better capture different data distributions and overlapping between clusters. These modifications were demonstrated to enhance the interpretability and robustness of fuzzy clustering in practical applications [5].

M. Menard et al. have proposed approaches to fuzzy clustering devising objective functions based on the principle of Extreme Physical Information. In their work, authors have exhibited how this method can systematically incorporate into the objective functions effectively minimal constraint terms. Their work is very well semantically explained for the physical perspective compared to the traditional algorithms [6].

H. Timm and R. Kruse have addressed a drawback in classical possibilistic fuzzy clustering, where the objective function is really minimized only if all cluster centers are equivalent and have proposed a modification to the objective function to trigger mutual repulsion among clusters, thus enhancing the clustering process [7].

J. Kang et al. have proposed a modified fuzzy c-means (FCM) algorithm that augments spatial neighborhood information into the conventional objective function. This approach was demonstrated to enhance the robustness of fuzzy clustering, especially when applied for image segmentation purposes [8].

H. Wang et al. have presented an automated multiscale fuzzy c-means (MSFCM) method for magnetic resonance images (MRI). The objective function of the conventional FCM method is modified to allow multiscale classification processing, thus improving the robustness especially when operating on low-contrast MR images [9].

X. Xiong et al. have proposed a modified generalized objective function for prototype-based fuzzy clustering incorporating a p-norm distance measure. Their approach induces cluster merging and the key innovation is the integration of principal component analysis (PCA) into the objective

function. This methodology successfully captures the directional structure of the clusters utilizing the principal components [10].
K. Zhao et al. have introduced a generalized fuzzy c-means (FCM) clustering strategy that modifies the objective function involving a mechanism to control the degree of fuzziness in clustering results. Via this this mechanism, the algorithm can tune between hard and fuzzy clustering, making it more adaptable to various datasets [11].
A. Bagherinia et al. have presented a reliability-driven cluster indicator to assess the reliability of fuzzy clusters within an ensemble framework. This methodology assigns weights to multiple clustering outcomes based on their reliability, which achieved an overall higher clustering quality and robustness [12].
The approach presented in this paper revives partially ideas by F. Hoppner et al. and K. Zhao et al., but the execution strategy is vastly different as instead of Lagrange multipliers, the numerical root mean squared propagation (RMSprop) method is employed.

3. The classical Fuzzy C-Means (FCM) algorithm

The classical Fuzzy C-Means algorithm (FCM) is the most significant algorithm in the field of fuzzy cluster analysis. It generalizes the well-known K-Means algorithm, involving the concept of partial degree of membership, thus allowing the instances of dataset to belong to several clusters simultaneously, with different degrees of membership. Therefore, the outcome of this algorithm, instead of being a set of clusters with their exclusive member points (like in the case of K-Means), will be the fuzzy membership matrix (U) containing the respective membership values of the instances into the clusters. The FCM algorithm operates as in iterative procedure aiming to approximate the nonlinear optimization of the objective function formulated classically as:

$$J(X,\ U,\ V;m,\ C)=\sum_{i=1}^{n}\sum_{j=1}^{C}\mu_{ij}^{m}d^{2}(x_i,\ c_j) \quad (1)$$

The hyperparameter m is the fuzzy exponent, which controls the degree of fuzziness in the generated clusters, i.e. the larger the value of m, the more distributed may be the instances into the clusters. Parameter C represents the number of clusters and it may either given externally as a hyperparameter, or tuned based on cluster validation measures. Additionally, n is the number of instances in the dataset, x_i is the i-th instance for $1 \leq i \leq n$, c_j is the centre of the jth cluster for $1 \leq j \leq C$ (i.e. the entries of the vector V) and μ_{ij} is the entry of membership matrix U corresponding to the ith element and the jth cluster. Furthermore, two other hyperparameters are the tolerance level Tol and the distance norm (typically the Euclidean distance) [13].
The algorithm runs in iterations where it updates the membership degree of each point into each cluster, then adjusts the cluster centers based on the memberships. Eventually the convergence is achieved, settling into a configuration that best fits the dataset. The hyperparameters collectively influence the clustering results, so good choices of them typically assisted by tuning procedures are helpful. The following pseudocode describes the FCM algorithm [14]:

1. Randomly initialize the centers of the clusters.
2. Initialize the fuzzy membership matrix with zero values.
3. Let k = 1 (iteration counter)
4. Evaluate the distances of data points from cluster centers (d_{ij} values).

5. Update the fuzzy membership matrix, according to: $\mu_{ij}=\frac{d_{ij}^{-\frac{2}{\varphi-1}}}{\sum_{k=1}^{c} d_{ik}^{-\frac{2}{\varphi-1}}}$
6. Calculate the new centers of the clusters, according to: $c_i=\frac{\sum_{j=1}^{n}\mu_{ij}^{\varphi}X_j}{\sum_{j=1}^{n}\mu_{ij}^{\varphi}}$
7. k = k+1 (increment the iteration counter)
8. If $\left\|U_{k-1}-U_{k-2}\right\|>Tol$ repeat at step 4.
9. END.

Despite the multiple successful applications of the FCM algorithm across a wide range of domains, it struggles when operating in datasets characterized by complex structures such as overlapping clusters, varying cluster sizes and shapes, presence of noise and or outliers. In such datasets, the FCM algorithm typically underperform resulting in poorly constructed clusters [15]. In this work, is proposed a modification (detailed in the next section) which modifies the objective function of FCM incorporating the notion of influence of a data point in the cluster. Moreover, the optimization approach in this work will be based on a numerical optimization method: the root mean squared method.

4. An entropy-based modified objective function FCM

The key idea of the modification proposed in this work is to control the influence that each point will have in the clustering process based on a reliability score that will be dynamically adapted. The definition of the objective function is:

$$J(X,\ U,\ V;m,\ C)=\sum_{i=1}^{n}\sum_{j=1}^{C} r_i\ \mu_{ij}^{m} d^2(x_i,\ c_j)+\lambda R(r) \tag{2}$$

In the above definition, X, U, V, m, C, μ_{ij}, c_j are the same as in the classical FCM, while r_i is the reliability score of each instance which will be dynamically updated, R(r) is a regularization term upon the vector of reliabilities and λ is a coefficient controlling the weight of the regularization term. The usage of the regularization term intends to avoid degenerate or extreme reliability assignments, inducing stability in the clustering process. There are several possibilities how the regularization term can be designated; in our case is used the L_2 regularization, which discourages extreme values in the reliability scores vector:

$$R(r)=\sum_{j=1}^{C} r_i^2 \tag{3}$$

On the other hand, in order to evaluate the reliability scores of the data points, an entropy-based approach is employed. For each data point will be evaluated the entropy value which quantifies how well distributed are the point memberships into the clusters:

$$H_i=-\sum_{j=1}^{C}\mu_{ij}\log\mu_{ij} \tag{4}$$

A high value of entropy indicates that a point is ambiguously participating in many clusters, while a low value of entropy (ideally zero) indicates a strong, reliable association of a point with a certain cluster. So, the value of the entropy can vary from 0 to $\log C$, where the value 0 indicates the highest reliability and the value $\log C$ indicates the lowest reliability. In the light of these facts, the evaluation of the reliability scores of the data points is handled as:

$$r_i = 1 - \frac{H_i}{\log C} \tag{5}$$

Obviously, the reliability scores of the points will vary from 0 (the lowest reliability) to 1 (the highest reliability).

In order to approximate a solution to the non-linear optimization problem defined by the given objective function, the root mean square propagation (RMSprop) method will be used. This method is an improvement over the classical gradient descent optimization technique. The fundamental principle of gradient descent is tracking the direction of the steepest descent, using a constant learning rate, while RMSprop adapts the learning rate adjusting it to the current landscape of the objective function. So generally, for a parameter θ, the update is handled as [16]:

$$\theta^{(k+1)} = \theta^{(k)} - \frac{\eta}{\sqrt{E[g^2]^{(k)} + \epsilon}} \cdot g^{(k)} \tag{6}$$

Parameter η represents the learning rate, $g^{(t)}$ denotes the gradient at the k-th iteration, $E[g^2]^{(k)}$ denotes the root mean square of the recent gradients and ε is a small constant to avoid division by zero. Furthermore, the update of the root mean square will be handled based on a decay parameter β as [17, 18]:

$$E[g^2_{\mu_{ij}}]^{(k+1)} = \beta \cdot E[g^2_{\mu_{ij}}]^{(k)} + (1-\beta) \cdot \left(g^{(t)}_{\mu_{ij}}\right)^2 \tag{7}$$

The general iterative scheme of the modified fuzzy clustering algorithm will remain the same as in the classical FCM, with the primary distinction that the updates for the cluster centres and the fuzzy membership values will be carried out numerically, according to the equations (8) and (9):

$$g^{(k)}_{c_j} = \frac{\partial J}{\partial c_j} = 2\left(\sum_{i=1}^{N} r_i \cdot \mu_{ij}^m \cdot (x_i - c_j)\right) + \lambda \cdot \frac{\partial R}{\partial c_{ij}} \tag{8}$$

$$g^{(k)}_{\mu_{ij}} = \frac{\partial J}{\partial \mu_{ij}} = r_i \cdot m \cdot \mu_{ij}^{m-1} d^2(x_i, c_j) + \lambda \cdot \frac{\partial R}{\partial \mu_{ij}} \tag{9}$$

Finally, the entire adaptive reliability-based fuzzy clustering algorithm is described by the following pseudocode:

1. Randomly initialize the centers of the clusters.
2. Initialize the fuzzy membership matrix with zero values.
3. Let k = 1 (iteration counter)
4. Update the gradients of the memberships, as: $g^{(k)}_{\mu_{ij}} = r_i \cdot m \cdot \mu_{ij}^{m-1} d^2(x_i, c_j) + \lambda \cdot \frac{\partial R}{\partial \mu_{ij}}$
5. Update the weighted root mean square, as:
$$E[g^2_{\mu_{ij}}]^{(k+1)} = \beta \cdot E[g^2_{\mu_{ij}}]^{(k)} + (1-\beta) \cdot \left(g^{(k)}_{\mu_{ij}}\right)^2$$

6. Update the fuzzy membership matrix: $\mu_{ij}^{(k+1)} = \mu_{ij}^{k} - \frac{\eta}{\sqrt{E[g_{\mu_{ij}}^2]^{(t)} + \epsilon}} \cdot g_{\mu_{ij}}^{(k)}$
7. Normalize the fuzzy membership matrix: $\mu_{ij}^{(k+1)} = \frac{\mu_{ij}^{(k+1)}}{\sum \mu_{ij}^{(k+1)}}$
8. Update the gradients of the centers, as: $g_{c_j}^{(k)} = 2\left(\sum_{i=1}^{N} r_i \cdot \mu_{ij}^{m} \cdot (x_i - c_j)\right) + \lambda \cdot \frac{\partial R}{\partial c_{ij}}$
9. Update the centers, as $c_j^{(k+1)} = c_j^{k} - \frac{\eta}{\sqrt{E[g_{c_{ij}}^2]^{(t)} + \epsilon}} \cdot g_{c_j}^{(k)}$
10. Update the reliability values: $r_i = 1 - \frac{H_i}{\log C}$
11. k = k+1 (increment the iteration counter)
12. If $\|U_k - U_{k-1}\| > Tol$ jump to step 4.
13. END.

5. Experimental results

In order to assess the robustness and the quality of the generated clusters generated by the proposed modified version of FCM algorithm, a series of experimental tests are conducted on several benchmark datasets. In order to evaluate the stability of the algorithm, in addition to the original versions of the benchmark datasets, two distorted versions are created for each benchmark dataset by adding artificial noise at different levels. The employed benchmark datasets were: Breast Cancer, Ionosphere, Vertebral Column, Dermatology, E. coli and Shuttle [19]. For each of the aforementioned datasets, two distorted versions are also created, with an additional quantity of respectively 2% and 5% noise points being added. The noise points are randomly placed at a distance from the cluster centres that is 8-10% larger than the average distance of the top-5 farthest genuine points from the respective cluster centre. The details of the original datasets are displayed in Table 1 below:

Table 1 - Summary of original benchmark datasets

Dataset	Number of attributes	Number of instances	Number of clusters
Breast Cancer	9	286	2
Ionosphere	34	351	2
Vertebral Column	6	310	3
Dermatology	34	366	6
E. coli	8	336	7
Shuttle	9	58000	7

Although the class labels for the employed classes are known, this information is not provided to the clustering procedures, instead it is utilized as the ground truth for the assessment of the clustering results. The quality assessment of the generated clusters is done via the V-measure, which is evaluated as the harmonic mean of homogeneity and completeness scores, so:

$$V = \frac{2}{\frac{1}{H} + \frac{1}{C}} \tag{10}$$

The homogeneity score measures the degree to which each cluster contains data points of only one particular label, while the completeness score measures how well all data points with the same label are grouped into the same cluster. In order to be compatible with the fuzzy scenario, firstly the conditional fuzzy entropy between the generated clusters and the ground truth is evaluated, and afterwards these results are utilized to calculate the fuzzy homogeneity and fuzzy completeness scores.
The results of the experimental procedures are summarized in Table 2 in the following page:

Table 2 – Fuzzy V-measure scores of the algorithms across the datasets

Dataset	Version	Classical FCM	Modified FCM
Breast Cancer	Original	0.82	0.83
	Distorted level 1	0.76	0.80
	Distorted level 2	0.67	0.75
Ionosphere	Original	0.74	0.76
	Distorted level 1	0.71	0.74
	Distorted level 2	0.65	0.70
Vertebral Column	Original	0.79	0.81
	Distorted level 1	0.73	0.77
	Distorted level 2	0.68	0.73
Dermatology	Original	0.70	0.73
	Distorted level 1	0.62	0.69
	Distorted level 2	0.57	0.67
Ecoli	Original	0.66	0.68
	Distorted level 1	0.60	0.65
	Distorted level 2	0.51	0.59
Shutttle	Original	0.62	0.64
	Distorted level 1	0.58	0.62
	Distorted level 2	0.52	0.59

As noticed from the table, the modified version of the FCM performs typically with a higher V-measure, pointing out the effectiveness of this approach. Moreover, it can be noticed that the difference in the V-measure values increases as the distortion level increases, which indicates a better robustness of this methodology. However, a drawback of this method is the increased computational complexity compared to the classical FCM.

6. Conclusions

This paper presented a modified reliability-based fuzzy clustering algorithm devised by modifications on the objective function of the classical FCM algorithm. The primary goal of this approach was to construct a more robust clustering framework, less sensitive to the presence of noise, outliers and uncertain instances. The proposed approach leverages entropy-based metrics to dynamically evaluate a confidence value for each data point, in order to control their influence during the clustering process. The optimization of the modified objective function is carried out by the RMSprop optimization techniques, in order to achieve a more flexible and adaptive clustering process.

The experimental procedures applied on several benchmark datasets and their slightly distorted variants demonstrated that the proposed algorithm generally performs with better fuzzy V-measure scores compared to the classical FCM algorithm. These findings indicate the method's improved resilience to noise and uncertain data and its capability to distinguish inherent clusters in challenging scenarios. Despite the natural computational overhead introduced by the adaptive modification, the overall gains in clustering quality and stability suggest that this reliability-based framework offers promising directions for robust fuzzy clustering in real-world applications.

Declaration on Generative AI

The author(s) have not employed any Generative AI tools.

References

[1] Oyewole, Gbeminiyi John, and George Alex Thopil. "Data clustering: application and trends." *Artificial intelligence review* 56, no. 7 (2023): 6439-6475.

[2] Ruspini, Enrique H., James C. Bezdek, and James M. Keller. "Fuzzy clustering: A historical perspective." *IEEE Computational Intelligence Magazine* 14, no. 1 (2019): 45-55.

[3] Lu, Jie, Guangzhi Ma, and Guangquan Zhang. "Fuzzy machine learning: A comprehensive framework and systematic review." *IEEE Transactions on Fuzzy Systems* (2024).

[4] Rada, Rexhep, Erind Bedalli, Sokol Shurdhi, and Betim Çiço. "A comparative analysis on prototype-based clustering methods." In *2023 12th Mediterranean Conference on Embedded Computing (MECO)*, pp. 1-5. IEEE, 2023.

[5] Höppner, Frank, Frank Klawonn, Rudolf Kruse, and Thomas Runkler. *Fuzzy cluster analysis: methods for classification, data analysis and image recognition.* John Wiley & Sons (1999): 71-78.

[6] Menard, Michel, and Michel Eboueya. "Extreme physical information and objective function in fuzzy clustering." *Fuzzy Sets and Systems* 128, no. 3 (2002): 285-303.

[7] Timm, Heiko, and Rudolf Kruse. "A modification to improve possibilistic fuzzy cluster analysis." In *2002 IEEE World Congress on Computational Intelligence. IEEE International Conference on Fuzzy Systems. FUZZ-IEEE'02. Proceedings*, vol. 2 (2002): 1460-1465

[8] Kang, Jiayin, Lequan Min, Qingxian Luan, Xiao Li, and Jinzhu Liu. "Novel modified fuzzy c-means algorithm with applications." *Digital signal processing* 19, no. 2 (2009): 309-319.

[9] Wang, Hesheng, and Baowei Fei. "A modified fuzzy C-means classification method using a multiscale diffusion filtering scheme." *Medical image analysis* 13, no. 2 (2009): 193-202.
[10] Xiong, Xuejian, Kap Chan, and Kian Lee Tan. "Similarity-driven cluster merging method for unsupervised fuzzy clustering." *arXiv preprint arXiv:1207.4155* (2012).
[11] Zhao, Kaixin, Yaping Dai, Zhiyang Jia, and Ye Ji. "General fuzzy c-means clustering strategy: Using objective function to control fuzziness of clustering results." *IEEE Transactions on Fuzzy Systems* 30, no. 9 (2021): 3601-3616.
[12] Bagherinia, Ali, Behrooz Minaei-Bidgoli, Mehdi Hosseinzadeh, and Hamid Parvin. "Reliability-based fuzzy clustering ensemble." *Fuzzy Sets and Systems* 413 (2021): 1-28.
[13] Bedalli, Erind, and Ilia Ninka. "Exploring an educational system's data through fuzzy cluster analysis." In *11th Annual International Conference on Information Technology & Computer Science*. 2014.
[14] Bedalli, Erind, Enea Mançellari, and Ozcan Asilkan. "A heterogeneous cluster ensemble model for improving the stability of fuzzy cluster analysis." *Procedia Computer Science* 102 (2016): 129-136.
[15] Liu, Zhe, and Sukumar Letchmunan. "Enhanced fuzzy clustering for incomplete instance with evidence combination." *ACM Transactions on Knowledge Discovery from Data* 18, no. 3 (2024): 1-20.
[16] Zou, Fangyu, Li Shen, Zequn Jie, Weizhong Zhang, and Wei Liu. "A sufficient condition for convergences of adam and rmsprop." In *Proceedings of the IEEE/CVF Conference on computer vision and pattern recognition*, pp. 11127-11135. 2019.
[17] Liu, Jinlan, Dongpo Xu, Huisheng Zhang, and Danilo Mandic. "On hyper-parameter selection for guaranteed convergence of RMSProp." *Cognitive Neurodynamics* 18, no. 6 (2024): 3227-3237.
[18] Babu, D. Vijendra, C. Karthikeyan, and Abhishek Kumar. "Performance analysis of cost and accuracy for whale swarm and RMSprop optimizer." In *IOP Conference Series: Materials Science and Engineering*, vol. 993, no. 1, p. 012080. IOP Publishing, 2020.
[19] Amarnath, B., S. Balamurugan, and Appavu Alias. "Review on feature selection techniques and its impact for effective data classification using UCI machine learning repository dataset." *Journal of Engineering Science and Technology* 11, no. 11 (2016): 1639-1646.