

Article

Enhancement of the Classification Performance of Fuzzy C-Means through Uncertainty Reduction with Cloud Model Interpolation

Weiwei Mao¹ and Kaijie Xu^{2,*}

¹ School of Intelligent Science and Information Engineering, Xi'an Peihua University, Xi'an 710125, China; 003603@peihua.edu.cn

² School of Electronic Engineering, Xidian University, Xi'an 710071, China

* Correspondence: kjsxu@xidian.edu.cn

Abstract: As an information granulation technology, clustering plays a pivotal role in unsupervised learning, serving as a fundamental cornerstone for various data mining techniques. The effective and accurate classification of data is a central focus for numerous researchers. For a dataset, we assert that the classification performance of a clustering method is significantly influenced by uncertain data, particularly those situated at the cluster boundaries. It is evident that uncertain data encapsulate richer information compared with others. Generally, the greater the uncertainty, the more information the data holds. Therefore, conducting a comprehensive analysis of this particular subset of data carries substantial significance. This study presents an approach to characterize data distribution properties using fuzzy clustering and defines the boundary and non-boundary characteristics (certainty and uncertainty) of the data. To improve the classification performance, the strategy focuses on reducing the uncertainty associated with boundary data. The proposed scheme involves inserting data points with the cloud computing technology based on the distribution characteristics of the membership functions to diminish the uncertainty of uncertain data. Building upon this, the contribution of boundary data is reassigned to the prototype in order to diminish the proportion of uncertain data. Subsequently, the classifier is optimized through data label (classification error) supervision. Ultimately, the objective is to leverage clustering algorithms for classification, thereby enhancing overall classification accuracy. Experimental results substantiate the effectiveness of the proposed scheme.



Citation: Mao, W.; Xu, K.

Enhancement of the Classification Performance of Fuzzy C-Means through Uncertainty Reduction with Cloud Model Interpolation.

Mathematics **2024**, *12*, 975. <https://doi.org/10.3390/math12070975>

Academic Editor: Xibei Yang

Received: 4 March 2024

Revised: 21 March 2024

Accepted: 24 March 2024

Published: 25 March 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: information granulation; fuzzy c-means (FCM); uncertain data; interpolation function; partition matrix; prototypes

MSC: 68T30; 68T42

1. Introduction

Clustering, fuzzy clustering (soft computing), and classification are fundamental techniques in data analysis and pattern recognition [1,2]. Clustering involves grouping similar data points, providing insights into inherent structures within datasets [3]. Fuzzy clustering, a subset of soft computing, extends traditional clustering methods by introducing the concept of membership function used for quantifying the extent to which data points belong to multiple clusters, thus capturing the inherent ambiguity present in many real-world scenarios [4]. In recent years, clustering has been recognized as an information granulation technology, and the clustering process is referred to as a granulation mechanism. This technology has found application in various fields. Classification, on the other hand, assigns predefined labels to data points based on their characteristics [5].

As we delve into the realm of uncertainty in data analysis, it becomes crucial to explore the significance of uncertain data. Unlike certain data, uncertain data encapsulate a greater volume of information due to their inherent variability and imprecision [6].

Understanding and effectively analyzing uncertainty in data have become essential in addressing the complexities of real-world applications [7]. In recent decades, a plethora of fuzzy set-based approaches [8] has emerged to effectively model the inherent uncertainty (granularity) present in various real-world phenomena. These methodologies aim to quantify information granularity by employing membership functions [9]. Among the many excellent algorithms in fuzzy clustering, fuzzy c-means (FCM) stands out as a popularly embraced soft partitioning algorithm, and is extensively applied across diverse domains. FCM plays a pivotal role in partitioning a given input space into distinct regions (groups, categories) based on a predefined similarity/dissimilarity measure [10]. Within the FCM algorithm, the dataset's underlying structure is articulated through partition matrices and prototypes (clusters) [11].

FCM utilizes membership function to measure the extent to which each data point (pattern) belongs to various clusters. From its origin, this technique has garnered considerable attention due to its application studies and conceptual developments [12]. It has proven to significantly enhance the quality of clustering and classification compared with the traditional hard partitioning methods. A plethora of enhanced clustering approaches have been developed over time. Among these alternatives, kernel-based FCM (KFCM) [13–15] has risen as an intriguing and widely adopted approach. KFCM employs various kernel functions that magically create nonlinear transformations, mapping the data from its native space into a higher-dimensional feature space. In this new space, the data are expected to exhibit greater separability. By positioning the data in this augmented space, KFCM aims to achieve superior classification performance [16]. The implicit nonlinear transformations provided by the kernel function contribute to capturing complex relationships within the data, potentially leading to more accurate and nuanced cluster assignments.

In summary, the KFCM algorithm extends the traditional FCM method by incorporating a kernel function. The key aspects to consider in the analysis of KFCM include the integration of the kernel function, its impact on nonlinear transformations and data separability in higher-dimensional spaces, and the resulting improvement in classification performance compared with conventional algorithms. A comprehensive analysis involves examining the specific kernel function employed, evaluating how nonlinear mappings contribute to enhanced data separability, and assessing the algorithm's robustness to varying data characteristics [17]. Additionally, investigating the computational complexity introduced by the kernel-based approach and exploring practical applications and use cases provide valuable insights into the algorithm's theoretical foundations and real-world effectiveness.

This research focused on the intersection of fuzzy clustering and uncertainty analysis. We aimed to investigate the meaningful application of fuzzy clustering in the context of uncertainty, leveraging its ability to model and accommodate imprecision and ambiguity in data. By incorporating fuzzy clustering into uncertainty analysis, we anticipated gaining valuable insights into the nuanced patterns and relationships within uncertain datasets, thus enhancing our ability to make informed decisions in the face of complexity and variability.

In the design process, we divided the data into certain and uncertain data based on the membership functions of each datum to all prototypes, and then inserted data with the cloud computing technology [18] based on the membership functions to reduce the proportion of uncertain data. Subsequently, reallocating the contribution of boundary data to the prototype was achieved to reduce the proportion of uncertain data. In this optimization process, we employed the classification error (data label) to supervise the insertion of data. Ultimately, the classification performance was improved by leveraging the enhanced partition matrix.

This paper is structured as follows: Section 2 provides a brief review of FCM and KFCM methods. In Section 3, we elucidate the principle behind the proposed scheme. Section 4 details the experimental studies over synthetic and publicly available data. Finally, Section 5 summarizes the study.

2. A Brief Overview of Fuzzy Set-Based Clustering

2.1. FCM Clustering Algorithm

FCM clustering comprises two primary phases: the computation of prototypes (cluster centers) and assigning points to the prototypes by utilizing a variation of Euclidean distance.

This iterative process continues until the prototypes reach a state of stabilization [19]. The method assigns membership values to data items for the clusters within a range of 0 to 1 by incorporating a fuzzification factor dictating the level of fuzziness within the clusters. The method aims to minimize the following objective function [20]

$$J_{FCM} = \sum_{i=1}^N \sum_{j=1}^C \mu_{ij}^m \|x_i - v_j\|^2$$

$$\mu_{ij} \geq 0, \sum_{j=1}^C \mu_{ij} = 1, \quad 0 < \sum_{j=1}^C \mu_{ij} < N$$
(1)

where x_i represents the i th data point from the dataset $X(X \in R^n)$, v_j stands for the cluster's j th prototype, μ_{ij} denotes the membership grade of the individual data point x_i that belongs to v_j , and $m(m > 1)$ is a scalar used to represent the fuzzification factor (coefficient) [21]. The fuzzification factor m plays a crucial role in shaping the clusters formed during the process. The symbol $\|\bullet\|$ denotes a distance function. The aforementioned objective function presented is minimized by iteratively calculating updates to the membership degree and prototypes [22], namely,

$$\mu_{ij} = \frac{1}{\sum_{k=1}^C \left(\frac{\|x_i - v_j\|}{\|x_i - v_k\|} \right)^{\frac{2}{m-1}}}$$
(2)

$$v_j = \frac{\sum_{i=1}^N (x_i \mu_{ij}^m)}{\sum_{i=1}^N \mu_{ij}^m}$$
(3)

2.2. Kernel-Based Fuzzy Clustering Approaches

Kernel-based fuzzy clustering (KFCM) is an algorithm that introduces kernel techniques into fuzzy clustering, aiming to overcome the limitations of traditional fuzzy clustering in handling nonlinear structures and high-dimensional data [23]. This algorithm incorporates a kernel function into the FCM framework, employing a nonlinear transformation to map data into a higher-dimensional feature space to better capture complex data structures.

One of the key innovations of the KFCM algorithm is the introduction of a kernel function. With the use of this function in fuzzy clustering, one can achieve a nonlinear transformation between the original data with a higher-dimensional feature space. This contributes to improving the separability of data, making clustering more effective in the new space. Through the nonlinear mapping provided by the kernel function, KFCM can flexibly handle nonlinear structures and intricate data relationships. Compared with traditional linear fuzzy clustering, this makes KFCM more suitable for datasets exhibiting nonlinear features. The KFCM algorithm constructs a fuzzy partition by defining the concept of membership grades of samples to each cluster. This consideration of fuzziness enhances the algorithm's tolerance to data uncertainty, facilitating a more comprehensive capture of data characteristics. KFCM demonstrates excellent nonlinear mapping capabilities in high-dimensional datasets. This advantage becomes particularly pronounced when dealing with datasets that have a large number of features. Similar to many other kernel methods, the selection of the kernel function and different parameter tuning will have significant influence on the performance of KFCM. When applying the KFCM algorithm, researchers

often need to carefully select a kernel function that suits the characteristics of the data and adjust parameters to achieve optimal performance.

In conclusion, KFCM provides a more flexible and powerful tool for fuzzy clustering by introducing kernel techniques, enabling better adaptation to nonlinear structures and high-dimensional data. In practical applications, researchers need to carefully balance the choice of the kernel function and parameter tuning to ensure the algorithm performs optimally for specific tasks. The objective function of the KFCM-F is mathematically designed in Equation (4) [24,25], and the associated constraints are the same as those in the FCM approach, as specified in (1).

$$J_{KFCM-F} = \sum_{i=1}^N \sum_{j=1}^C u_{ij}^m \|\Phi(x_i) - \Phi(v_j)\|^2 \tag{4}$$

The key merit of the KFCM-F lies in the placement of prototypes within the feature space, which are magically mapped to a space with higher-dimensional features by the utilization of a suitable kernel function [26]. Then, Equation (4) can be computed in the following way:

$$\begin{aligned} \|\Phi(x_i) - \Phi(v_j)\|^2 &= \Phi(x_i)^T \Phi(x_i) - 2\Phi(x_i)^T \Phi(v_j) + \Phi(v_j)^T \Phi(v_j) \\ &= K(x_i, x_i) - 2K(x_i, v_j) + K(v_j, v_j) \end{aligned} \tag{5}$$

With the use of the technique of Lagrange multipliers, one can execute the optimizing of the membership degree and prototypes, and can obtain the form below:

$$\mu_{ij} = \frac{1}{\sum_{k=1}^C \left[\frac{1-K(x_i, v_j)}{1-K(x_i, v_k)} \right]^{\frac{1}{m-1}}} \tag{6}$$

$$v_j = \frac{\sum_{i=1}^N \mu_{ij}^m K(x_i, v_j) x_i}{\sum_{i=1}^N \mu_{ij}^m K(x_i, v_j)} \tag{7}$$

The algorithm for KFCM-F closely resembles the original form of FCM. It initiates with a random membership matrix and iteratively updates with the prototypes until a specified stopping criterion is met.

3. Enhancing Fuzzy Clustering through Innovative Interpolation Techniques

As previously highlighted, when being used to deal with classification tasks, an algorithm’s performance is particularly influenced by the presence of uncertain data situated at the clusters’ boundaries. This part of the data affects the position of the prototype, which in turn affects the partition matrix and ultimately affects the algorithm’s classification performance. Therefore, our focus is on optimizing these cluster boundaries.

In fuzzy clustering, the focus is traditionally placed on the maximum values within each column of the membership matrix, as these values play a crucial role in determining the clustering (classification) results. However, we should recognize that non-maximal values also carry valuable information. For instance, these non-maximal values can provide insights into data points situated at the boundaries of clusters, contributing to a more comprehensive understanding of the clustering structure. While the maximum values heavily influence the overall outcome, acknowledging the significance of non-maximal values enhances the nuanced interpretation of the clustering results.

Let μ_{imaxj} represent the j th largest value of $\mu_i (\mu_i = [\mu_{i1}, \mu_{i2}, \dots, \mu_{iC}])$. We use the standard deviation of membership values to partition the data into two parts, namely certain (non-boundary) data X_c and uncertain (boundary) data X_u .

To enhance the optimization of cluster boundaries, we aim to decrease the proportion of data within these boundaries (considered uncertain data). This can be effectively accomplished by increasing the proportion of certain data. Consequently, we incorporate additional data into the set of certain data. The method we adopted in this study is to insert data X_I with the cloud computing technology into the certain data according to the membership functions, and modify the prototype matrix based on the new dataset. During the interpolation process, we used the one-dimensional normal membership cloud model to determine the distribution of the interpolated data of each feature.

The cloud model serves as an uncertainty conversion framework that translates a specific qualitative concept expressed through natural language values. It primarily comprises both the forward and backward cloud generators. In this study, our focus lies on data generation for interpolation, primarily driven by the principles underlying the forward cloud generators.

The forward cloud generator functions as a mapping tool that translates qualitative information into quantitative data by utilizing three numerical characteristic parameters of the cloud, namely, expectation (Ex), entropy (En), and super-entropy (He), along with the count of cloud droplets (N). The output of this process provides the quantitative positioning of N cloud droplets within the numerical field space, accompanied by the confidence level associated with each droplet, representing the underlying concept. Given the widespread applicability of normal clouds, the screening process primarily revolves around their utilization.

A specific method for the one-dimensional forward cloud generator is outlined as follows:

Input: Digital parameters (Ex, En, He) embodying the quality concept of weight and the enumeration of cloud droplets (N).

Output: N cloud droplets (X_i) and the association of each cloud droplet with the conceptual framework.

- a. In the initial phase, a normal random number (En_i) is produced, where the expectation is set to En and the standard deviation to He .
- b. Subsequently, a normally distributed random number (X) is generated with Ex as the mean and En_i as the standard deviation.
- c. In the third phase, the degree of X 's association with the specified concept is determined through:

$$u_i = \exp \left[\frac{(x - Ex)^2}{2En_i^2} \right] \quad (8)$$

The process is iteratively repeated through steps *a* to *c* until N cloud droplets have been successfully generated.

During algorithm execution, the determination of the actual parameters of cloud models is subject to the prototype and the standard deviation of the membership degrees. Then, the prototypes can be adjusted based on the new dataset \bar{X} according to (3):

$$\bar{v}_j = \frac{\sum_{i=1}^N (\bar{x}_i \mu_{ij}^m)}{\sum_{i=1}^N \mu_{ij}^m} \quad (9)$$

Afterward, the partition matrix can undergo additional refinement utilizing the adjusted prototype matrix. During this optimization stage, we utilize the classification error (original labels of the dataset) to supervise the interpolation process. This refinement of the partition matrix ultimately leads to an enhancement in classification performance. Figure 1 illustrates the methodology of implementing the proposed scheme in detail.

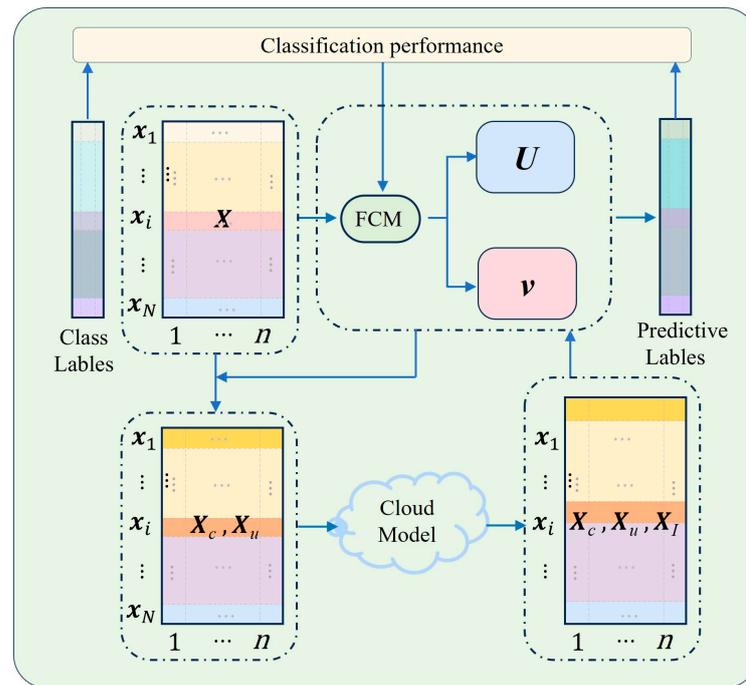


Figure 1. An overall model: main functional processing phases.

The total computational complexity of KFCM is $O(CN^2n)$, while for FCM and the proposed method the computational complexities are respectively $O(CNn)$ and $O(C(N + N_I)n)$ (C is the number of prototypes, N is the total number of the original data instances in the n -dimensional space, and N_I stands for the number of inserted data in the n -dimensional space). Typically, the number of uncertain data (N_I) is much smaller than the total number of data (N). Therefore, theoretically, our algorithm’s execution speed will be slower than the FCM but much faster than the KFCM.

4. Experimental Studies

In what follows, we aimed to assess the effectiveness of the developed scheme by comparing its performance with that of the FCM and Gaussian kernel function-based FCM (KFCM-G) methods. The primary goal of this extensive series of experiments was to discuss the classification performance of these clustering approaches. A variety of experiments were conducted using both synthetic datasets and publicly available datasets (<http://archive.ics.uci.edu/ml>, accessed on 3 March 2024) [27].

All data were normalized with the min-max scaling method, which is described as follows:

$$x' = \frac{x - \min(\mathbf{X})}{\max(\mathbf{X}) - \min(\mathbf{X})} \tag{10}$$

where x and x' represent the original and the preprocessed data values, respectively. The goal was to thoroughly evaluate the proposed scheme’s effectiveness. To ensure consistency, all data were normalized to $[0, 1]$. The classification rate [4] was utilized as the primary metric in these experiments, given its widespread usage as an index for performance evaluation.

In the experiments, we explored various values being positioned in the interval $[1.1, 3.1]$ for the fuzzification factor, and changed its value with a step size of 0.2. The number of iterations was fixed at 500 to ensure the completion of clustering. We permitted the methods to terminate if the following condition was met:

$$\max(\|\mathbf{u} - \mathbf{u}'\|) \leq 10^{-5} \tag{11}$$

where \mathbf{U}' represents the membership matrix coming from the previous iteration. In numerous instances, Equation (10) was fulfilled before the maximum iteration had been reached. We let the Gaussian kernel parameter σ^2 vary from 10 to 100 in increments of 10 to mitigate the computational intensity associated with KFCM(G).

To gauge the efficacy of the proposed approach, we employed 10-fold cross-validation [28], a widely utilized technique for estimating and validating the classification performance and stability of the fuzzy classification models.

4.1. Synthetic Data Experiments

The first experiment utilized a two-dimensional synthetic dataset comprising 450 individuals categorized into nine distinct classes. The dataset's geometry is illustrated in Figure 2. Figures 3–5 present the clustering outcomes along with the corresponding partition matrix of the three approaches. The experimental results associated with the classification rates and the model parameter values of the synthetic dataset are plotted in Figure 6. It was evident that through the allocation of a judiciously chosen fuzzy factor to each datum and the incorporation of the interpolation technology, the optimization of prototypes occurred. This resulted in the refinement of class boundaries.

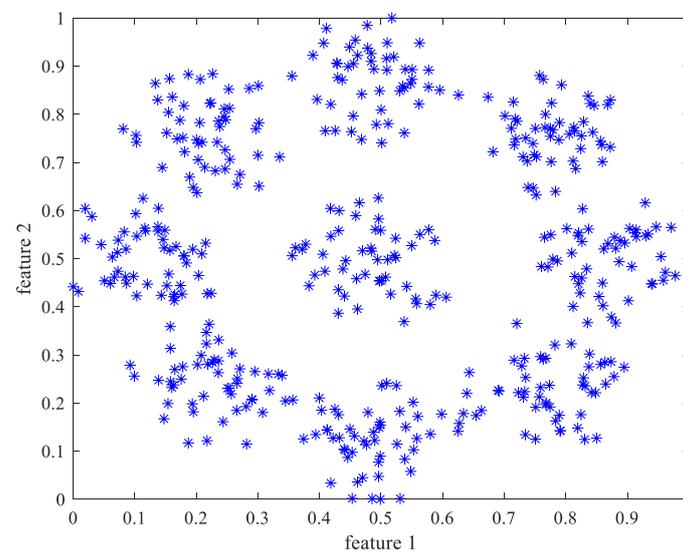


Figure 2. Synthetic dataset.

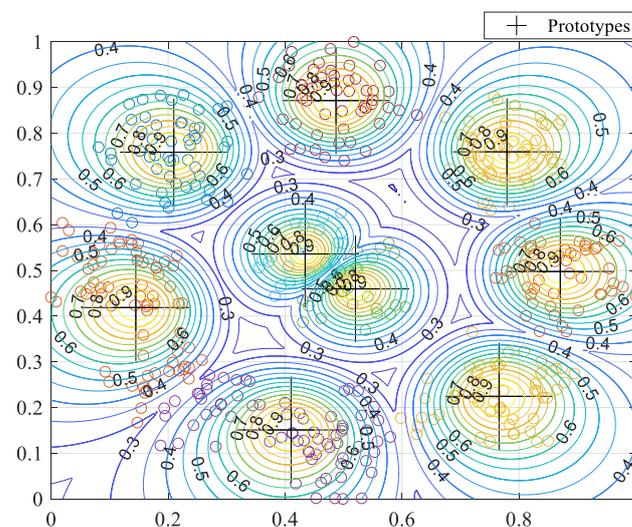


Figure 3. Clustering results of the synthetic dataset with FCM.

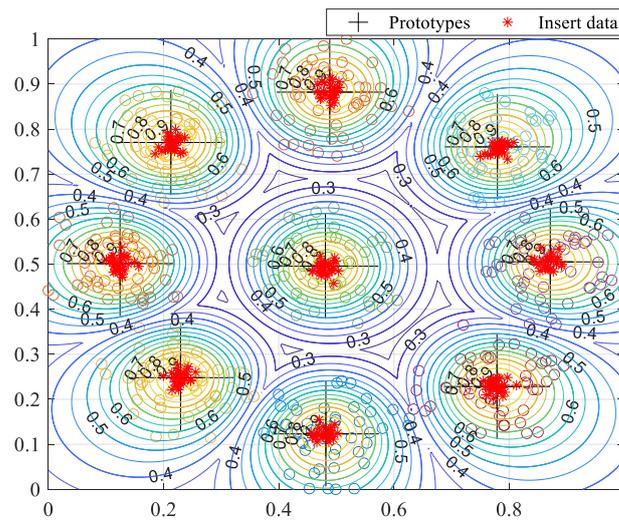


Figure 4. Clustering results of the synthetic dataset with the proposed method.

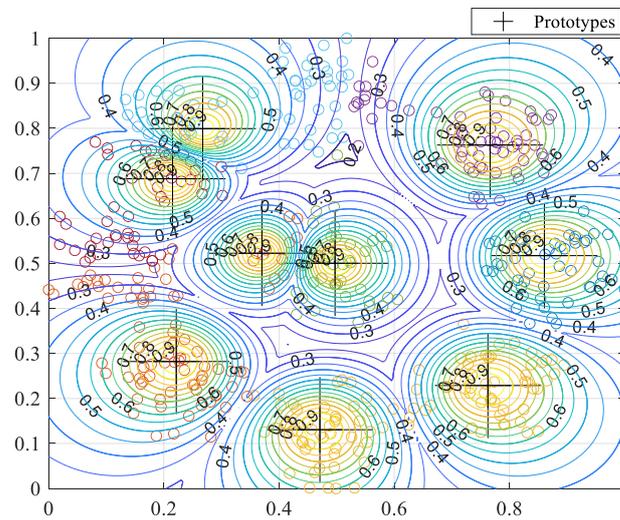


Figure 5. Clustering results of the synthetic dataset with KFCM.

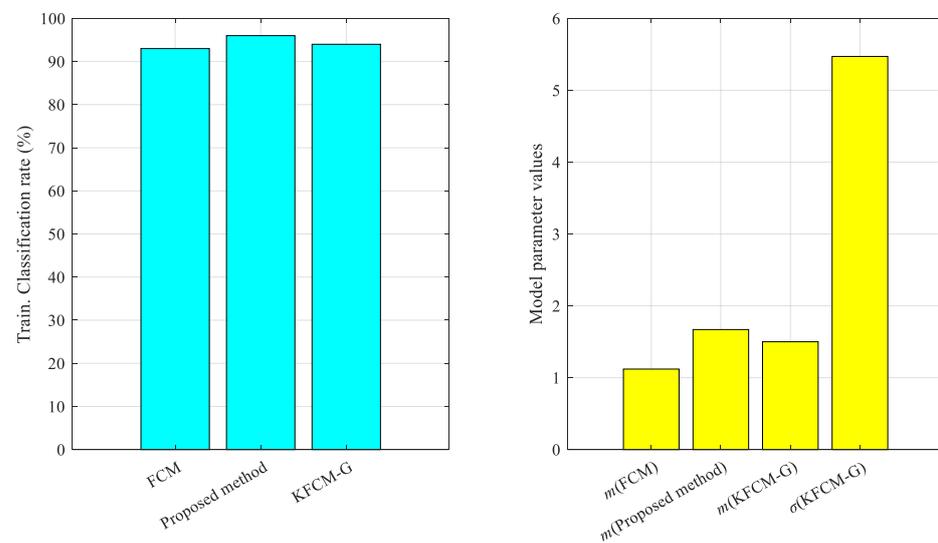


Figure 6. Classification rate results of the synthetic dataset.

4.2. Publicly Available Data Experiments

We employed six publicly accessible datasets, detailed descriptions of which are available in the UCI machine learning repository. Figures 7–12 show the experimental results associated to the classification rates and the model parameter values of each dataset. It is noteworthy that the classification quality of all these datasets was enhanced through the application of the proposed method. The developed scheme exhibited substantial merits over both FCM and KFCM methods.

The KFCM exhibited improvements on some specific datasets. Notably, the developed approach consistently achieved higher classification rates compared with both the FCM and other kernel-based clustering algorithms. This superiority can be attributed to the optimization of the proposed method on the cluster boundaries through incorporating specific data into the clustering process, thereby reducing the proportion of uncertain data and refining the prototypes. Consequently, this optimization facilitated more accurate cluster identification.

The observed enhancement in classification performance averaged approximately 6%, with improvements ranging from a minimum of 3% to a maximum of 10%. This range represents the most notable improvement achieved by our method.

In summary, our approach achieved the partitioning of uncertain and deterministic data by utilizing membership degrees to delineate boundary and non-boundary data. Based on this, we leveraged the cloud modeling technology for data interpolation and adjusted the prototype to refine the partition matrix, further enhancing the model's classification performance. This not only enriched and advanced classifier models based on fuzzy clustering technology but also offered valuable insights for uncertainty analysis research.

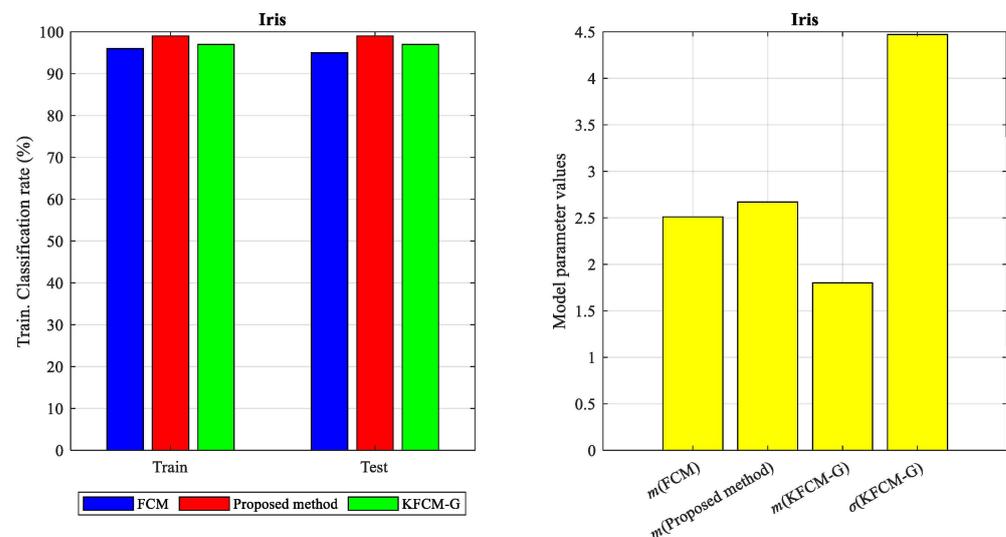


Figure 7. Classification rate results of the Iris dataset.

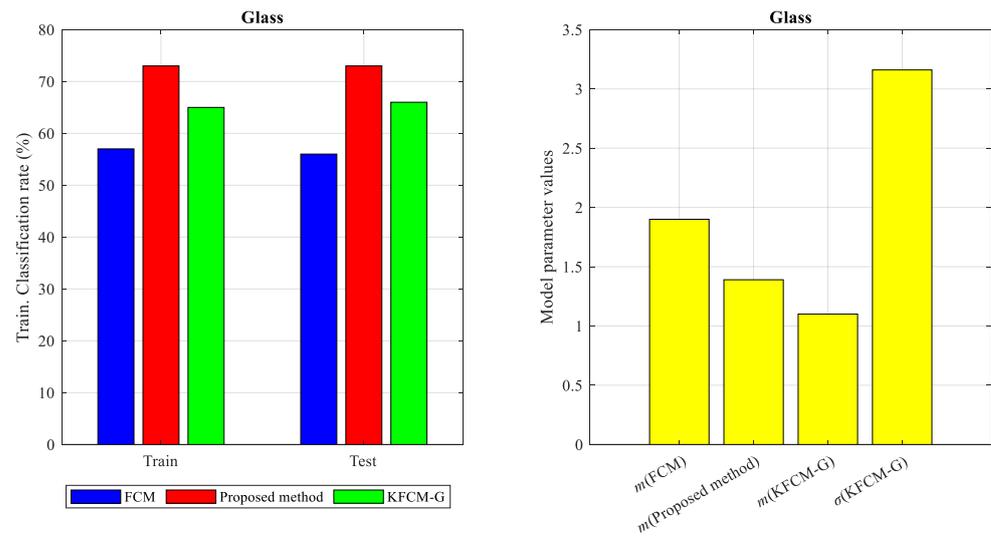


Figure 8. Classification rate results of the Glass dataset.

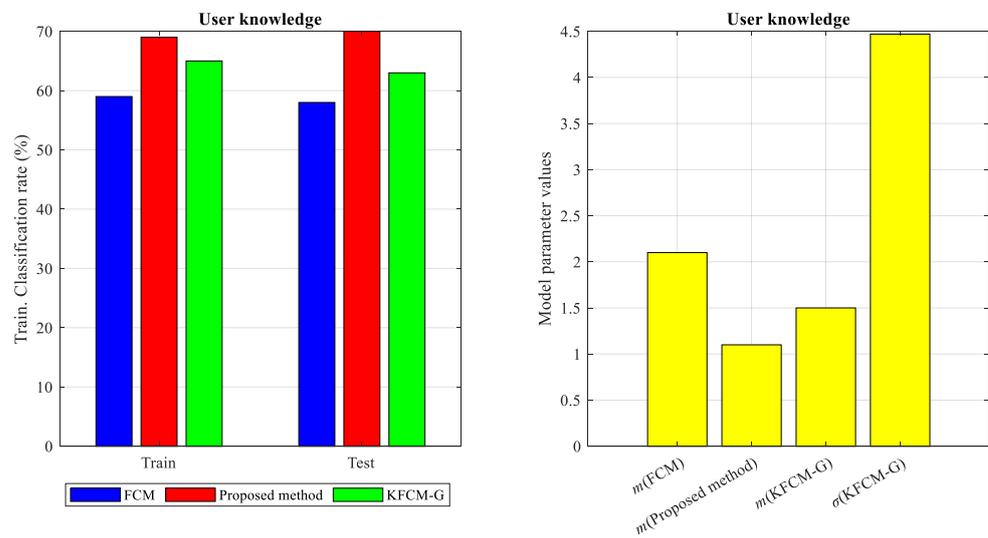


Figure 9. Classification rate results of the User Knowledge dataset.

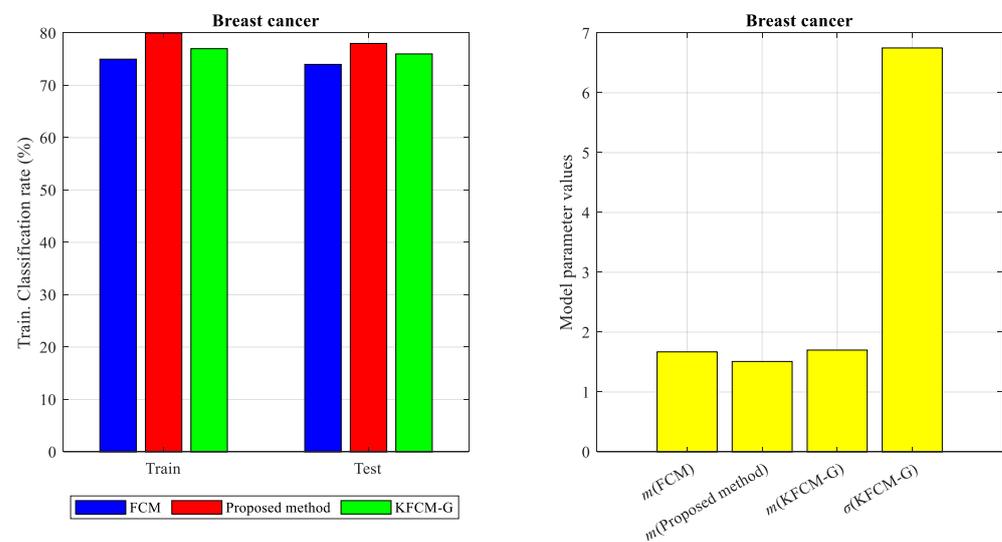


Figure 10. Classification rate results of the Breast Cancer dataset.

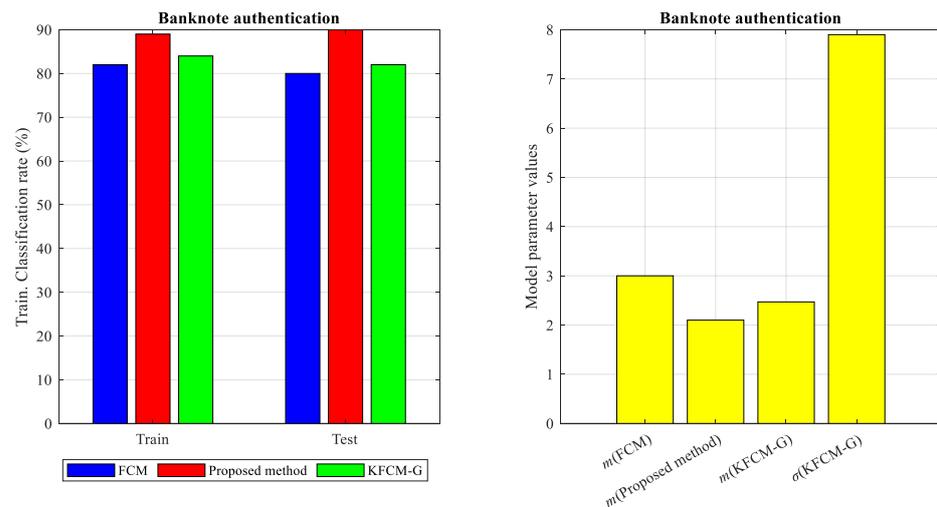


Figure 11. Classification rate results of the Banknote Authentication dataset.

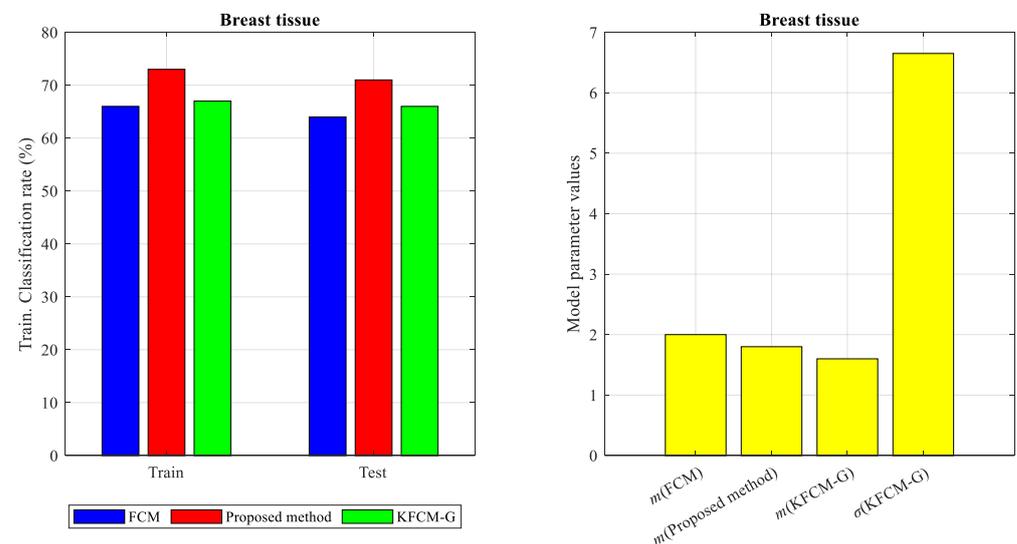


Figure 12. Classification rate results of the Breast Tissue dataset.

5. Conclusions

This study developed a new interpolation technique with cloud computing aimed at improving the efficacy of fuzzy clustering. In the design phase, we utilized the membership functions of each data point concerning all prototypes to distinguish between uncertain data (located at the cluster boundaries) and certain data. To optimize these cluster boundaries, we established a function to insert specific data into the certain data, thereby reducing the proportion of uncertain data. This new dataset enabled the modification of prototypes. The interpolation principle here was implemented using cloud models. Subsequently, the partition matrix of the original dataset was refined using these modified prototypes, ultimately improving the performance of fuzzy clustering.

The theoretical works provided in this study are underpinned by a series of experimental studies. The proposed method not only provides a unique avenue for improving the quality of fuzzy clustering but also raises broader questions about interpolation techniques.

Author Contributions: The authors confirm contribution to the paper as follows: study conception and design: K.X.; data collection: W.M.; analysis and interpretation of results: W.M.; draft manuscript preparation: W.M. All authors reviewed the results and approved the final version of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Natural Science Foundation of China (Nos. 62101400, 72101075, 72171069 and 92367206), in part by the China Postdoctoral Science Foundation under Grant 2023M732743, and in part by the Shaanxi Fundamental Science Research Project for Mathematics and Physics under Grant 22JSQ032.

Data Availability Statement: Data will be made available on request.

Acknowledgments: Yixi Wang (School of Electronic Engineering, Xidian University) made contributions during the first round of revision.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Wang, L.; Cui, G.; Cai, X. Fuzzy clustering optimal k selection method based on multi-objective optimization. *Soft Comput.* **2023**, *27*, 1289–1301. [[CrossRef](#)]
2. Sharma, H.K.; Kar, S. Decision making for hotel selection using rough set theory: A case study of Indian hotels. *Int. J. Appl. Eng. Res.* **2018**, *13*, 3988–3998.
3. Cerqueti, R.; Mattera, R. Fuzzy clustering of time series with time-varying memory. *Int. J. Approx. Reason.* **2023**, *153*, 193–218. [[CrossRef](#)]
4. Vovan, T.; Nguyenhoang, Y.; Danh, S. An automatic fuzzy clustering algorithm for discrete elements. *J. Oper. Res. Soc. China* **2023**, *11*, 309–325. [[CrossRef](#)]
5. Burés, J.; Larrosa, I. Organic reaction mechanism classification using machine learning. *Nature* **2023**, *613*, 689–695. [[CrossRef](#)]
6. Hamidzadeh, J.; Moradi, M. Enhancing data analysis: Uncertainty-resistance method for handling incomplete data. *Appl. Intell.* **2020**, *50*, 74–86. [[CrossRef](#)]
7. Antonopoulou, H.; Mamalougou, V.; Theodorakopoulos, L. The role of economic policy uncertainty in predicting stock return volatility in the banking industry: A big data analysis. *Emerg. Sci. J.* **2022**, *6*, 569–577. [[CrossRef](#)]
8. Das, A.K.; Chakraborty, B.; Goswami, S.; Chakrabarti, A. A fuzzy set based approach for effective feature selection. *Fuzzy Sets Syst.* **2022**, *449*, 187–206. [[CrossRef](#)]
9. Hanyu, E.; Cui, Y.; Pedrycz, W.; Li, Z. Design of Distributed Rule-Based Models in the Presence of Large Data. *IEEE Trans. Fuzzy Syst.* **2023**, *31*, 2479–2486.
10. Cui, Y.; Hanyu, E.; Pedrycz, W.; Li, Z. Augmentation of rule-based models with a granular quantification of results. *Soft Comput.* **2019**, *23*, 12745–12759. [[CrossRef](#)]
11. Hanyu, E.; Cui, Y.; Pedrycz, W.; Fayek, A.R.; Li, Z.; Li, J. Design of fuzzy rule-based models with fuzzy relational factorization. *Expert Syst. Appl.* **2022**, *206*, 117904.
12. Zhu, X.B.; Pedrycz, W.; Li, Z.W. Fuzzy clustering with nonlinearly transformed data. *Appl. Soft Comput.* **2017**, *61*, 364–376. [[CrossRef](#)]
13. Devijver, P.A.; Kittler, J. *Pattern Recognition: A Statistical Approach*; Prentice-Hall: London, UK, 1982.
14. Xu, K.J.; Pedrycz, W.; Li, Z.W.; Nie, W.K. High-accuracy signal subspace separation algorithm based on gaussian kernel. *IEEE Trans. Ind. Electron.* **2019**, *66*, 491–499. [[CrossRef](#)]
15. Bouchachia, A.; Pedrycz, W. Enhancement of fuzzy clustering by mechanisms of partial supervision. *Fuzzy Sets Syst.* **2006**, *157*, 1733–1759. [[CrossRef](#)]
16. Wang, J.H.; Lee, W.J.; Lee, S.J. A kernel-based fuzzy clustering algorithm. In Proceedings of the International Conference on Innovative Computing, Information and Control, Beijing, China, 30 August–1 September 2006; IEEE Computer Society: Washington, DC, USA, 2006; pp. 550–553.
17. Singh, S.; Srivastava, S. Kernel fuzzy C-means clustering with teaching learning based optimization algorithm (TLBO-KFCM). *J. Intell. Fuzzy Syst.* **2022**, *42*, 1051–1059. [[CrossRef](#)]
18. Osadcha, K.P.; Osadchyi, V. The use of cloud computing technology in professional training of future programmers. *CTE Workshop Proc.* **2021**, *8*, 155–164. [[CrossRef](#)]
19. Grover, N. A study of various fuzzy clustering algorithms. *Int. J. Eng. Res.* **2014**, *3*, 177–181. [[CrossRef](#)]
20. Hanyu, E.; Cui, Y.; Pedrycz, W.; Li, Z. Enhancements of rule-based models through refinements of Fuzzy C-Means. *Knowl.-Based Syst.* **2019**, *170*, 43–60.
21. Cui, Y.; Hanyu, E.; Pedrycz, W.; Li, Z. Designing Distributed Fuzzy Rule-Based Models. *IEEE Trans. Fuzzy Syst.* **2021**, *29*, 2047–2053. [[CrossRef](#)]
22. Dey, A.; Senapati, T.; Pal, M.; Chen, G. A novel approach to hesitant multi-fuzzy soft set based decision-making. *AIMS Math.* **2020**, *5*, 1985–2008. [[CrossRef](#)]
23. Huang, X.; Zhang, C.Z.; Yuan, J. Predicting extreme financial risks on imbalanced dataset: A combined kernel FCM and kernel SMOTE based SVM classifier. *Comput. Econ.* **2020**, *56*, 187–216. [[CrossRef](#)]
24. Shen, H.; Yang, J.; Wang, S.; Liu, X. Attribute weighted mercer kernel based fuzzy clustering algorithm for general non-spherical datasets. *Soft Comput.* **2006**, *10*, 1061–1073. [[CrossRef](#)]

25. Zhang, D.Q.; Chen, S.C. Clustering incomplete data using kernel-based fuzzy c-means algorithm. *Neural Process. Lett.* **2003**, *18*, 155–162. [[CrossRef](#)]
26. Graves, D.; Pedrycz, W. Kernel-based fuzzy clustering and fuzzy clustering: A comparative experimental study. *Fuzzy Sets Syst.* **2010**, *161*, 522–543. [[CrossRef](#)]
27. Elkano Ilintxeta, M.; Sanz Delgado, J.A.; Barrenechea Tartas, E.; Bustince Sola, H.; Galar Idoate, M. CFM-BD: A distributed rule induction algorithm for building compact fuzzy models in Big Data classification problems. *IEEE Trans. Fuzzy Syst.* **2020**, *28*, 163–177. [[CrossRef](#)]
28. Bavan, L.; Surmacz, K.; Beard, D. Adherence monitoring of rehabilitation exercise with inertial sensors: A clinical validation study. *Gait Posture* **2019**, *70*, 211–217. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.