

Article

Semi-Supervised Fuzzy Clustering Based on Prior Membership

Yinghan Hong¹, Guoxiang Zhong², Jiahao Lian^{3,*}, Guizhen Mai¹, Honghong Zhou⁴, Pinghua Chen³, Junliu Zhong¹ and Hui Cao¹

¹ School of Artificial Intelligence, Guangzhou Maritime University, Guangzhou 510520, China; hongyinghan@gzmtu.edu.cn (Y.H.); maiguizhen@gzmtu.edu.cn (G.M.); zhongjunliu@gzmtu.edu.cn (J.Z.)

² Pengcheng Laboratory, Shenzhen 518000, China

³ School of Computer, Guangdong University of Technology, Guangzhou 510006, China

⁴ Guangdong Science and Technology Innovation Monitoring and Research Center, Guangzhou 510030, China; zhouhhg@163.com

* Correspondence: 19128159280@163.com; Tel.: +86-19128159280

Abstract

Traditional fuzzy clustering algorithms construct sample partition criteria solely based on similarity measures but lack an effective representation of prior membership information, which limits further improvements in clustering accuracy. To address this issue, this paper proposes a semi-supervised fuzzy clustering algorithm based on prior membership (SFCM-PM). The proposed algorithm introduces prior information entropy as a metric to quantify the divergence between partition membership and prior membership and incorporates this as an auxiliary partition criterion into the objective function. By jointly optimizing data similarity and consistency with prior knowledge during the clustering process, the algorithm achieves more accurate and reliable clustering results. The experimental results demonstrate that the SFCM-PM algorithm achieves significant performance improvements by incorporating a small number of prior membership samples across several standard and real-world datasets. It also performs outstandingly on datasets with unbalanced sample distributions.

Keywords: clustering; prior membership; prior information entropy; partition criterion

MSC: 62H30



Academic Editor: Hsien-Chung Wu

Received: 9 July 2025

Revised: 7 August 2025

Accepted: 8 August 2025

Published: 10 August 2025

Citation: Hong, Y.; Zhong, G.; Lian, J.; Mai, G.; Zhou, H.; Chen, P.; Zhong, J.; Cao, H. Semi-Supervised Fuzzy Clustering Based on Prior Membership. *Mathematics* **2025**, *13*, 2559. <https://doi.org/10.3390/math13162559>

Copyright: © 2025 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/>).

1. Introduction

Clustering is an important research area in machine learning. Due to its effective applications in fields such as image processing [1–6] and recommendation systems [7–9], it is currently attracting increasing attention from researchers for in-depth study. Traditional clustering methods are typical unsupervised learning algorithms. Compared with classification, clustering does not rely on predefined classes and class labels for training samples; instead, it partitions samples into several clusters based on their similarity. However, traditional clustering algorithms cannot further improve the accuracy of sample partitioning in scenarios where prior membership information is available.

Traditional clustering methods can generally be divided into two categories: (1) hard clustering based on k-means clustering [10] and (2) soft clustering based on fuzzy c-means clustering (fuzzy c-means clustering, FCM) [11]. K-means clustering was first proposed by Anderberg. The algorithm initially generates several cluster centers randomly and then calculates the similarity between each sample and each cluster center, assigning the sample to the cluster with the highest similarity. Subsequently, the cluster centers are recalculated,

and the above steps are repeated until the partitioning results no longer change. This algorithm uses the Euclidean distance to measure similarity and strictly partitions samples into specific clusters. Based on this, the FCM algorithm considers the fuzziness of samples belonging to clusters, using membership degrees to represent the soft partitioning of samples. Its fundamental idea is to maximize the similarity among samples within the same cluster while minimizing the similarity among samples in different clusters. Both types of traditional unsupervised clustering algorithms fail to effectively model prior membership information. However, in practical applications, prior membership information can often be obtained through domain knowledge or expert input, which has significant potential to enhance clustering performance but remains underutilized.

Wagstaff et al. [12] proposed a semi-supervised clustering algorithm based on k-means clustering (constrained k-means algorithm). This algorithm strictly adheres to prior information during the sample partitioning process and has some effect on improving clustering accuracy, although it does not directly utilize prior membership degrees as the form of prior information. Yasunori et al. [13] proposed a semi-supervised fuzzy c-means clustering model (Semi-supervised c-means clustering, SFCM), which introduces prior membership into similarity measurements. Later, Yin et al. [14] proposed a semi-supervised metric-based fuzzy clustering algorithm based on the SFCM algorithm, which incorporates metric learning and entropy regularization. In fact, to make maximum use of prior membership information, it is necessary to establish a sample partitioning criterion that relates partitioning membership to prior membership [15].

In recent years, significant progress has been made in fuzzy clustering research that incorporates prior information. Zhao et al. [16] proposed a fuzzy c-means clustering strategy based on a generalized objective function (general fuzzy c-means clustering, GFCM), which enhances the flexibility and interpretability of the algorithm by adjusting the form of the objective function to control the fuzziness of clustering results. Wang et al. [17] employed entropy regularization to handle pairwise constraints but did not consider direct optimization of the membership correction term. To fully utilize prior membership information, Yao Lan et al. [18] proposed a robust cross-entropy fuzzy clustering algorithm (CEFCM), which optimizes clustering results by measuring membership differences through cross-entropy. Li Le et al. [19] adopted a hierarchical strategy to process prior information, facilitating the transfer of prior knowledge across different levels.

Although previous studies have made progress in incorporating prior information [20] into clustering, most do not explicitly model the relationship between partition membership and prior membership within the objective function. This limitation hinders the full utilization of available prior knowledge and constrains further improvements in clustering accuracy. Recent advances in semi-supervised clustering provide promising strategies to address this issue. For example, Gaussian mixture models (GMMs) [21] offer probabilistic sample assignments but rely on parametric assumptions that may not suit complex real-world data. Semi-supervised Gaussian process modeling [22] leverages kernel-based similarity measures to integrate prior knowledge, which inspires our use of prior information entropy. Furthermore, fuzzy decision-making frameworks [23] show that embedding prior constraints into the objective function can enhance optimization robustness—this insight directly informs the design of our prior information entropy term. In addition, studies in fuzzy random logistics networks [24] and stochastic power distribution systems [25] demonstrate that integrating prior constraints via fuzzy mathematics can improve solution robustness. These works motivate the development of a unified objective function for our proposed method, SFCM-PM. Moreover, entropy regularization techniques used in cross-condition fault diagnosis [26] highlight the effectiveness of entropy metrics in improving model consistency, which supports our choice to use prior information entropy for

measuring membership discrepancies. Therefore, this paper proposes a semi-supervised fuzzy clustering algorithm based on prior membership (SFCM-PM), aiming to improve clustering accuracy by effectively leveraging prior membership information. By introducing prior information entropy as a measure of divergence between partition membership and prior membership, our method enhances the clustering process through a unified optimization framework.

The main contributions of this work are summarized as follows:

- We propose a new semi-supervised fuzzy clustering algorithm named SFCM-PM, which introduces prior information entropy as an auxiliary partition criterion to quantify the divergence between partition membership and prior membership.
- The algorithm integrates sample similarity and prior consistency into a unified objective function, allowing effective utilization of limited prior information.
- Extensive experiments are conducted on standard and real-world datasets, demonstrating the superior performance of SFCM-PM over traditional and semi-supervised clustering algorithms, especially in imbalanced data scenarios.

The remainder of this paper is organized as follows. Section 2 reviews related work, Section 3 presents our methodology, Section 4 details the experimental validation, Section 5 analyzes the experimental results, and Section 6 concludes this paper.

2. Related Work

Semi-supervised fuzzy clustering algorithms vary in design but all aim to achieve effective clustering by combining auxiliary information with the fuzziness of the data. Next, we will introduce the fuzzy c-means algorithm closely related to this study.

Fuzzy c-Means Clustering

The fuzzy c-means clustering (FCM) [11] algorithm partitions a given dataset $X = \{x_1, x_2, \dots, x_n\}$ (where n is the number of samples) into c clusters by minimizing an objective function based on the squared error between data points and cluster centers, weighted by membership degrees. The objective function is defined as follows:

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

$$\text{s.t.} \quad \sum_{i=1}^c \mu_{ij} = 1 \quad (2)$$

where $U = [\mu_{ij}]$ is the membership matrix, and μ_{ij} represents the degree to which sample belongs to the i -th cluster. $V = [v_1, v_2, \dots, v_c]$ is the cluster center matrix, and $v_i \in V$ denotes the center of the i -th cluster. m is the fuzziness coefficient, with $m = 2$ generally considered optimal. The partitioning criterion of fuzzy c-means clustering is the minimum sum of squared errors, i.e., minimizing $J(U, V; X)$ under constrained conditions. Therefore, the update formulas for the membership matrix $U = [\mu_{ij}]$ and cluster center $V = [v_1, v_2, \dots, v_c]$ are as follows, respectively:

$$\mu_{ij} = \frac{[d^2(x_j, v_i)]^{\frac{1}{1-m}}}{\sum_{i=1}^c [d^2(x_j, v_i)]^{\frac{1}{1-m}}} \quad (3)$$

$$v_i = \frac{\sum_{j=1}^n \mu_{ij}^m x_j}{\sum_{j=1}^n \mu_{ij}^m} \quad (4)$$

This paper addresses these challenges through three key contributions:

- Proposal of prior information entropy as a direct metric for quantifying membership divergence.
- Development of a dual-criteria optimization framework that combines data similarity and prior consistency.
- Theoretical analysis of convergence properties with mathematical guarantees.

3. Methodology

This section presents the proposed semi-supervised fuzzy clustering algorithm based on prior membership (SFCM-PM). We begin by elaborating on the core concept and motivation behind the algorithm. This is followed by a detailed description of the algorithm itself, including the design of the novel objective function that incorporates prior information entropy. Subsequently, the optimization solution method used to derive the update rules for cluster centers and membership degrees is presented. The complete algorithmic steps are then outlined. Finally, a complexity analysis is provided to evaluate the computational efficiency of the proposed approach.

3.1. Semi-Supervised Fuzzy Clustering Based on Prior Membership

3.1.1. Basic Idea of the Algorithm

In fuzzy clustering, membership degrees are used to quantify the degree of sample affiliation with each cluster. To facilitate distinction, we refer to the membership obtained through fuzzy clustering as partition membership and the membership derived from prior knowledge as prior membership. Compared to partition membership, prior membership is generally closer to the true affiliation of samples. Therefore, during the clustering process, it is essential to minimize the discrepancy between partition membership and prior membership to enhance the accuracy and reliability of clustering results. This idea constitutes the core concept of the proposed semi-supervised fuzzy clustering approach based on prior membership. To this end, this paper proposes a semi-supervised fuzzy clustering algorithm based on prior membership (SFCM-PM). The proposed algorithm introduces prior information entropy as a measurement criterion to evaluate the difference between partition membership and prior membership, and it incorporates this as an auxiliary partitioning criterion into the objective function. During optimization, the algorithm simultaneously considers both data similarity and consistency with prior knowledge, thereby achieving more accurate clustering results under limited prior information guidance.

The SFCM-PM algorithm is designed with practical scenarios in mind, where prior information may not always be fully reliable. The method exhibits inherent robustness to both incomplete and inaccurate prior information:

- **Handling Incomplete Information:** The algorithm is fundamentally a semi-supervised approach. The penalty coefficient ω_j in the objective function (Equation (6)) acts as a switch: $\omega_j = 1$ for samples with available prior membership and $\omega_j = 0$ for unlabeled samples. This means that the term prior information entropy only influences the clustering process for the supervised subset. For the vast majority of unlabeled data, the algorithm reverts to a data-driven approach, relying solely on sample similarity. This design ensures that the algorithm remains effective even when only a small fraction of the data is labeled, which is the common case in real-world applications.
- **Robustness to Inaccurate Information:** The use of prior information entropy as a soft constraint, rather than a hard constraint, is key to the robustness of the algorithm. Unlike methods that strictly enforce prior labels (e.g., must-link/cannot-link constraints), our entropy term gently guides the membership degrees towards the prior. If a prior label is incorrect, the algorithm can still correct it through joint optimization of the objective function, which balances the prior consistency term ($\sum_{j=1}^n \omega_j H(\bar{\mu}_j, \mu_j)$) with

the data similarity term $(\sum_{i=1}^c \sum_{j=1}^n (\mu_{ij} - \bar{\mu}_{ij})^m d^2(x_j, v_i))$. The fuzziness coefficient m further provides flexibility, allowing memberships to adjust even for labeled samples. This prevents the algorithm from being overly sensitive to noisy or erroneous prior labels.

In terms of practicality, the SFCM-PM algorithm is well-suited for scenarios where a small amount of high-quality prior knowledge can be obtained and significantly influences clustering outcomes. Typical applications include water quality assessment based on water color imagery and environmental quality monitoring, among others. In water quality assessment using water color images, leveraging a limited set of expert-labeled samples to guide the clustering of large volumes of similar images can substantially improve the accuracy of pollution-level classification. For regional air quality evaluation involving multiple pollutants, clustering historical sensor data with the aid of a small quality-verified labeled dataset helps identify pollution patterns and anomalous events, thereby facilitating dynamic monitoring and early warning of environmental conditions.

The SFCM-PM algorithm is not dependent on fully reliable prior clustering. Its design allows it to gracefully handle sparse and potentially noisy prior information, making it a practical and robust tool for real-world semi-supervised clustering tasks.

3.1.2. Algorithm Description

The SFCM-PM algorithm constructs an objective function that integrates both sample similarity measures and prior consistency constraints by introducing prior information entropy as a metric. Based on this objective function, update rules for the cluster centers and membership matrix are derived. The algorithm adopts an alternating optimization strategy, updating the cluster centers and membership matrix iteratively until convergence is achieved. In the SFCM-PM algorithm, the difference between the two is characterized by defining prior information entropy, as shown in Definition 1 below.

Definition 1. For a sample $x_j \in X$, if the prior membership is $\bar{\mu}_j = \{\bar{\mu}_{1j}, \bar{\mu}_{2j}, \dots, \bar{\mu}_{cj}\}$ (where c is the number of clusters) and the partition membership is $\mu_j = \{\mu_{1j}, \mu_{2j}, \dots, \mu_{cj}\}$, prior information entropy is defined as follows:

$$\begin{aligned} H(\bar{\mu}_j, \mu_j) &= - \sum_{i=1}^c \bar{\mu}_{ij} \ln \mu_{ij} \\ \text{s.t. } \sum_{i=1}^c \mu_{ij} &= 1 \end{aligned} \quad (5)$$

The incorporation of prior information entropy as an auxiliary partition criterion is grounded in information theory and offers several key advantages over conventional approaches for modeling prior knowledge in semi-supervised clustering.

The primary motivation stems from the need to directly minimize the divergence between the evolving partition membership μ_j and the fixed prior membership $\bar{\mu}_j$. In information theory, entropy is a fundamental measure of uncertainty or surprise. By defining the prior information entropy as $H(\bar{\mu}_j, \mu_j) = - \sum_{i=1}^c \bar{\mu}_{ij} \ln \mu_{ij}$, the algorithm effectively quantifies the “surprise” or “cost” associated with the current partition membership μ_j given the prior knowledge $\bar{\mu}_j$. Minimizing this cost in the objective function drives the partition membership μ_j to become more consistent with prior membership $\bar{\mu}_j$, thereby leveraging the available supervision.

The main reasons why we chose this specific form of entropy are as follows:

1. **Direct Optimization:** Unlike methods that use prior information to modify distance metrics (e.g., SFCM [13]) or impose hard constraints (e.g., constrained k-means [10]),

our entropy term directly optimizes the membership values themselves. This allows for a more precise and targeted adjustment of the fuzzy partition.

2. **Soft and Probabilistic Guidance:** The entropy term acts as a soft regularization. It does not force a sample to belong to a specific cluster (hard constraint) but gently encourages the membership distribution to align with the prior. This is particularly beneficial when prior information might be noisy or incomplete, as it prevents overfitting to potentially erroneous labels while still guiding the clustering process.
3. **Mathematical Tractability:** The entropy function is smooth and differentiable with respect to the membership degree μ_{ij} . This property is crucial for deriving the update rules via gradient-based optimization, ensuring a stable and efficient iterative solution.

Compared to Euclidean distance measures, entropy more effectively captures the differences in information content between probability distributions. Unlike KL divergence, our formulation demonstrates greater robustness when partition memberships approach zero. Furthermore, in contrast to constraint-based methods that impose hard restrictions, our approach incorporates a soft penalty, enabling more flexible and stable optimization.

In essence, prior information entropy provides a theoretically sound, mathematically elegant, and practically effective mechanism to seamlessly integrate prior knowledge into the fuzzy clustering process, directly guiding the formation of the membership matrix towards a solution that is both data-driven and prior-informed.

3.1.3. Objective Function Design

During sample partitioning, the SFCM-PM algorithm uses not only sample feature-based similarity as the primary partitioning criterion but also incorporates sample membership-based differences as an auxiliary criterion. The objective function is defined as follows:

$$J_{\text{SFCM-PM}}(U, V; X, \bar{U}) = \sum_{i=1}^c \sum_{j=1}^n (\mu_{ij} - \bar{\mu}_{ij})^m d^2(x_j, v_i) + \sum_{j=1}^n \omega_j H(\bar{\mu}_j, \mu_j) \quad (6)$$

where $\bar{U} = [\bar{\mu}_{ij}]$ is the prior membership matrix, and ω_j is the penalty coefficient for sample x_j . For sample $x_j \in X$, if $\bar{\mu}_j = \mathbf{0}$, this indicates that no prior membership is available for x_j , and $\omega_j = 0$; if $\bar{\mu}_j \neq \mathbf{0}$, prior membership is available, and $\omega_j = 1$.

The fundamental idea of the SFCM-PM algorithm is to not only maximize the similarity within clusters and minimize similarity between clusters but also minimize the difference between partition memberships and prior memberships. Therefore, it is necessary to minimize $J_{\text{SFCM-PM}}(U, V; X, \bar{U})$.

3.1.4. Optimization Solution Method

To minimize the objective function $J_{\text{SFCM-PM}}(U, V; X, \bar{U})$ under the normalization constraint, the Lagrangian function for $J_{\text{SFCM-PM}}(U, V; X, \bar{U})$ is constructed based on the constraint in Definition 1:

$$\begin{aligned} L_{\text{SFCM-PM}}(U, V; X, \bar{U}, \lambda) = & \sum_{i=1}^c \sum_{j=1}^n (\mu_{ij} - \bar{\mu}_{ij})^m d^2(x_j, v_i) + \sum_{j=1}^n \omega_j H(\bar{\mu}_j, \mu_j) \\ & + \sum_{j=1}^n \lambda_j \left(\sum_{i=1}^c \mu_{ij} - 1 \right) \end{aligned} \quad (7)$$

Taking the partial derivative of $L_{\text{SFCM-PM}}(U, V; X, \bar{U}, \lambda)$ with respect to V and setting it to zero yields the following:

$$\left. \frac{\partial L_{SFCM-PM}(U, V; X, \bar{U}, \lambda)}{\partial V} \right|_{V=V^*} = 0 \quad (8)$$

Thus, the update formula for cluster centers is

$$v_i = \frac{\sum_{j=1}^n (\mu_{ij} - \bar{\mu}_{ij})^m x_j}{\sum_{j=1}^n (\mu_{ij} - \bar{\mu}_{ij})^m} \quad (9)$$

Taking the partial derivative of $L_{SFCM-PM}(U, V; X, \bar{U}, \lambda)$ with respect to U and setting it to zero yields

$$\left. \frac{\partial L_{SFCM-PM}(U, V; X, \bar{U}, \lambda)}{\partial U} \right|_{U=U^*} = 0 \quad (10)$$

That is,

$$m(\mu_{ij} - \bar{\mu}_{ij})^{m-1} d^2(x_j, v_i) - \omega_j \frac{\bar{\mu}_{ij}}{\mu_{ij}} + \lambda_j = 0 \quad (11)$$

Given $\sum_{i=1}^c \mu_{ij} = 1$, the update formula for sample memberships can be further derived as follows:

$$\mu_{ij} = \bar{\mu}_{ij} + \left(1 - \sum_{i=1}^c \bar{\mu}_{ij} \right) \frac{[md^2(x_j, v_i) + \omega_j(\mu_{ij} - \bar{\mu}_{ij})^{1-m}]^{\frac{1}{1-m}}}{\sum_{i=1}^c [md^2(x_j, v_i) + \omega_j(\mu_{ij} - \bar{\mu}_{ij})^{1-m}]^{\frac{1}{1-m}}} \quad (12)$$

When $\bar{U} = \mathbf{0}$, the cluster center and membership update formulas of the SFCM-PM algorithm are consistent with those of the FCM algorithm, indicating that the SFCM-PM algorithm is an extension of the FCM algorithm in scenarios with prior membership.

In the above derivation, Lagrange multipliers are used to ensure the normalization property of the membership matrix, but they are not updated during actual iterative solving. Therefore, the role of Lagrange multipliers mainly manifests in the theoretical derivation stage and is not updated as independent variables in specific algorithm implementations.

3.2. Algorithm Steps

The SFCM-PM algorithm operates through an iterative process to find the optimal cluster center V and membership matrix U . The algorithm begins by initializing the membership matrix $U(0)$ based on the number of clusters c and the sample dataset. In each iteration, the algorithm performs two main update steps:

1. **Cluster Center Update:** The cluster center $V(t) = [v_1, v_2, \dots, v_c]$ is recalculated using the cluster center update formula (Equation (9)), which incorporates both the similarity of the data and the prior consistency. This step adjusts the cluster prototypes based on the weighted positions of the data points, where the weights are derived from the difference between the current partition memberships μ_{ij} and the prior memberships $\bar{\mu}_{ij}$.
2. **Membership Update:** The membership degrees μ_{ij} for each sample x_j belonging to each cluster i are updated using the membership update formula (Equation (11)). This update rule is influenced by three factors: the distance $d^2(x_j, v_i)$ between the sample and the center of the cluster, the prior membership $\bar{\mu}_{ij}$ for the sample-cluster pair, and the current membership values themselves (implicitly through the normalization term). This ensures that the new memberships reflect both the similarity of the data and the consistency with the prior information provided (controlled by ω_j).

Iterations continue until a stopping criterion is met, such as reaching the maximum number of iterations T or changes in the membership matrix U falling below a predefined

threshold ε . Upon convergence, the final membership matrix $U(t)$ is used to assign group label L' to the samples, typically selecting the group with the highest degree of membership for each sample. The algorithm flow chart is shown in Figure 1.

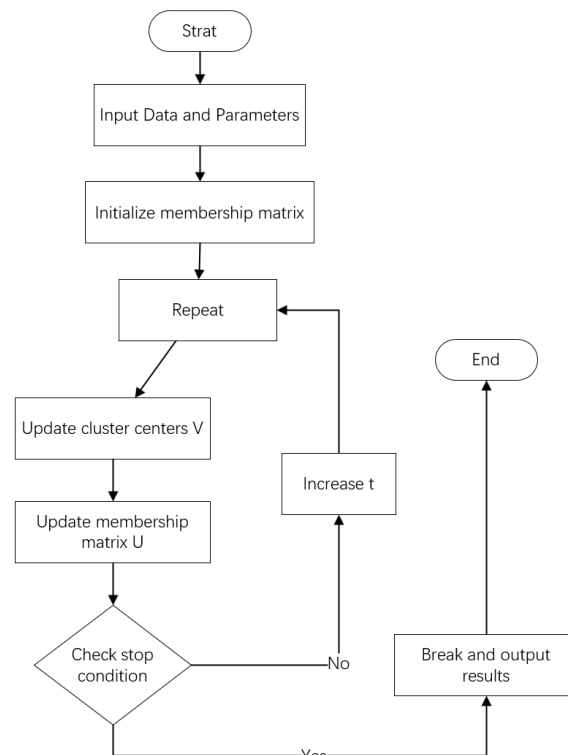


Figure 1. This flowchart illustrates the entire process of the SFCM-PM algorithm, which, under the condition of integrating prior membership degrees, iteratively optimizes the objective function and gradually converges to generate the final clustering labels.

Algorithm 1 outlines the steps of the SFCM-PM algorithm, which is based on its mathematical derivation and related formulas.

Algorithm 1 SFCM-PM algorithm parameters and process.

Input: Samples set $X = \{x_1, x_2, \dots, x_n\}$, Prior membership $\bar{U} = [\bar{\mu}_{ij}]$, Clusters number c , Allowable error ε , Max iterations T , Penalty coefficient ω_j .

Output: Cluster labels $L' = \{l'_1, l'_2, \dots, l'_n\}$

Process:

- 1 Initialize the membership $U(0) = [\mu_{ij}]$ by cluster number and samples set
 - 2 Repeat
 - 3 Calculate cluster centers $V(t) = [v_1, v_2, \dots, v_c]$
 - 4 Calculate the membership $U(t) = [\mu_{ij}]$
 - 5 If $t < T$ or $\|U(t) - U(t-1)\| > \varepsilon$
 - 6 break;
 - 7 Else
 - 8 $t = t + 1$
 - 9 End
 - 10 Return the membership $U(t)$ and then confirm the cluster label $L' = \{l'_1, l'_2, \dots, l'_n\}$
-

The table presents the complete specification of the SFCM-PM clustering algorithm, including input parameters, computational steps, and output format.

3.3. Algorithm Complexity Analysis

The workflow of the SFCM-PM algorithm primarily involves two core steps: updating cluster centers and sample memberships. Based on these steps, its time complexity is analyzed as follows.

For cluster centers, each iteration requires traversing all samples to calculate weighted sums and normalization factors according to cluster center update Formula (8), with a time complexity of $O(cdn)$, where d is the sample dimension. For sample memberships, membership update Formula (11) involves distance calculations, exponential operations, and normalization, with a time complexity of $O(cn(d + m))$, where m is the number of samples with prior membership. When the number of algorithm updates is t , the time complexity of the SFCM-PM algorithm is $O(tcn(d + m) + tcdn)$. In practical applications, $c \ll n$, $d \ll n$, and $m \ll n$ are typically satisfied. Therefore, the algorithm's time complexity $O(tcn(d + m) + tcdn)$ is linear. Although the SFCM-PM algorithm introduces slight additional computations by incorporating prior information entropy as an auxiliary partitioning criterion, it does not alter the overall complexity growth trend, demonstrating good scalability and practicality.

4. Experimental Section

This section describes experiments conducted on eight different datasets. For each dataset, 5%, 10%, 15%, and 20% of the samples are randomly selected to generate prior membership information. To verify the performance of the algorithm, classic clustering algorithms such as k-means, FCM, SFCM, SSFCM, and cGFCM were selected as comparison benchmarks.

4.1. Datasets

The experimental datasets were selected from commonly used datasets in the UCI and KEEL databases, including Segment, Digits-389, Parkinsons, Ecoli, Iris, and Banknote Authentication. These datasets are widely used for validating clustering algorithms. Segment, Digits-389, and Iris have balanced sample numbers across clusters, while Parkinsons, Ecoli, and Banknote Authentication have significant differences in sample numbers across clusters, and they are mainly used to validate the algorithm's performance on unbalanced datasets. Furthermore, to test the effectiveness of the algorithm in practical cases, the Air dataset was selected for regional air quality evaluation with multiple pollution factors, and the Water dataset was selected for water quality evaluation based on water color images from Reference [27]. Table 1 provides relevant information on these datasets.

Table 1. Dataset details with attribute count.

Dataset	Number of Samples	Number of Attributes	Number of Clusters
Segment	2310	19	7
Digits-389	3165	16	3
Parkinsons	195	22	2
Ecoli	336	7	8
Iris	150	4	3
Banknote Authentication	1372	4	2
Water	173	9	3
Air	284	6	3

For all datasets, the number of clusters specified for the algorithms matches the true number of classes, ensuring a fair and reasonable evaluation of clustering performance.

4.2. Baselines and Evaluation Metric

To comprehensively evaluate the effectiveness of the proposed SFCM-PM algorithm, five representative clustering methods were selected as baseline approaches:

- K-Means: The classic hard clustering algorithm iteratively updates the cluster centers by minimizing the squared error within the clusters. Each sample belongs only to a single cluster, and the algorithm does not have the capability to model prior information; it serves as a benchmark for unsupervised clustering [10].
- FCM: The foundational work in fuzzy clustering introduces membership degrees to describe the attribution probability of samples to clusters. Soft partitioning is achieved by minimizing the weighted squared error, but prior information is not utilized [11].
- SFCM: The semi-supervised fuzzy clustering algorithm converts pairwise constraints into membership degree constraints. The supervisory information was integrated by modifying the objective function, but the differences in prior membership degrees were not explicitly optimized [13].
- SSFCM: Based on the improved algorithm of SFCM, entropy regularization and a metric learning mechanism are introduced. The fuzziness of the membership degrees is controlled by regularization terms to enhance the robustness to noisy data, but the use of previous membership degrees is still insufficient [28].
- cGFCM: A general FCM clustering algorithm was developed based on Minkowski distance and contraction mapping. This algorithm updates the prototype by constructing a contraction mapping [16].

These benchmark algorithms cover a range from traditional unsupervised to semi-supervised fuzzy clustering models. Specifically, k-means and FCM are unsupervised and do not incorporate prior information, while SFCM, SSFCM, and cGFCM introduce various forms of prior constraints or regularization mechanisms to improve clustering performance. Comparing SFCM-PM against these baselines enables a thorough evaluation of its ability to model prior membership and achieve more accurate clustering under limited supervision.

The experiments use the commonly used Rand index (RI) [29–31] to evaluate the performance of the clustering algorithm. This metric ranges from $[0, 1]$, with values closer to 1 indicating better clustering performance and values closer to 0 indicating poorer performance. Suppose the dataset is $X = \{x_1, x_2, \dots, x_n\}$, the actual class labels are $L = \{l_1, l_2, \dots, l_n\}$, and the class labels obtained by the clustering algorithm are $L' = \{l'_1, l'_2, \dots, l'_n\}$. Considering pairwise sample comparisons, the following definitions apply:

$$\begin{aligned} TP &= \{(x_j, x_k) \mid l_j = l_k, l'_j = l'_k, j \neq k\} \\ FP &= \{(x_j, x_k) \mid l_j \neq l_k, l'_j = l'_k, j \neq k\} \\ TN &= \{(x_j, x_k) \mid l_j \neq l_k, l'_j \neq l'_k, j \neq k\} \\ FN &= \{(x_j, x_k) \mid l_j = l_k, l'_j \neq l'_k, j \neq k\} \end{aligned}$$

The definitions of the components are as follows:

- True Positive (TP): The number of sample pairs that are in the same cluster in both the ground truth and the clustering result.
- False Positive (FP): The number of sample pairs that are in the same cluster in the clustering result but in different clusters in the ground truth.
- True Negative (TN): The number of sample pairs that are in different clusters in both the ground truth and the clustering result.

- False Negative (FN): The number of sample pairs that are in different clusters in the clustering result but in the same cluster in the ground truth.

These metrics are used to compute the Rand Index, which measures the similarity between two clustering results. The Rand Index (RI) is defined as follows:

$$RI = \frac{|TP| + |TN|}{|TP| + |FP| + |FN| + |TN|} \quad (13)$$

To statistically compare the overall performance of multiple algorithms across datasets, the Friedman test is employed. For k algorithms tested on N datasets, let R_j be the average rank of the j -th algorithm. The Friedman statistic is calculated as follows:

$$\chi_F^2 = \frac{12N}{k(k+1)} \left[\sum_{j=1}^k R_j^2 - \frac{k(k+1)^2}{4} \right] \quad (14)$$

Under the null hypothesis (no performance difference between algorithms), χ_F^2 follows a χ^2 distribution with $k - 1$ degrees of freedom.

This experiment is based on the Rand Index (RI), and Friedman evaluated the performance of the algorithm.

4.3. Implementation Details

All experiments were completed in the unified computing environment shown in Table 2. The algorithm proposed in this paper was evaluated under the same conditions as the other algorithms while maintaining consistent parameter settings. Each group of experiments was repeated 50 times, and the average value was taken as the final result.

Table 2. Experimental platform information.

Structure	Configuration
CPU	Intel(R) Core(TM) i7-6700K @ 4.00 GHz 4.00 GHz
Memory	16.0 GB
Operating System	64bit Windows 10
Software Platform	PyCharm 2022.2.2

Prior Membership Generation Strategy

To rigorously evaluate the performance of the SFCM-PM algorithm under semi-supervised conditions, a systematic and reproducible procedure was employed to generate the prior membership information for a subset of data samples. The prior membership was constructed from the ground-truth class labels available in the benchmark datasets, simulating a realistic scenario where a small, randomly selected portion of the data is labeled by an expert or through an auxiliary labeling process. The generation process consisted of the following steps:

1. Random Sampling: For each dataset, a predefined proportion (5%, 10%, 15%, or 20%) of the total samples were randomly selected without replacement. This subset represents the “supervised” portion of the data.
2. Label Assignment: The ground-truth class label for each selected sample was used to define its prior membership. Specifically, a one-hot encoding scheme was applied: for a sample known to belong to cluster i , its prior membership vector was set to 1 for the i -th component and 0 for all other components.
3. Handling Unlabeled Data: For the remaining samples (the “unsupervised” portion), no prior membership information was assumed. This was mathematically represented by setting all components of their prior membership vector to zero.

This approach ensures that the prior information is both accurate (derived from ground truth) and sparse (available for only a small fraction of samples), providing a fair and challenging testbed to evaluate the algorithm's ability to leverage limited supervision. The random selection process was independently repeated for each of the 50 experimental runs to ensure the robustness of the reported average performance.

4.4. Experimental Result Analysis

4.4.1. Overall Performance Comparison

The key experimental parameters include the following: maximum iteration number $T = 200$, algorithm termination threshold $\varepsilon = 10^{-5}$, and fuzziness index $m = 2$. To ensure fairness, all data were standardized to the $[0, 1]$ interval before use to avoid the impact of different dimensions on distance calculations. Tables 3–7 show the performance comparison results of the SFCM-PM algorithm and other algorithms under different prior sample ratios.

In summary, the experimental results presented in Tables 3–7 consistently demonstrate the effectiveness of the proposed SFCM-PM algorithm. When no prior membership information is available (0% prior samples, Table 3), SFCM-PM performs comparably to the standard FCM algorithm, as expected, since the prior information entropy term becomes inactive. However, as the proportion of samples with previously provided membership increases (5% to 20%), SFCM-PM exhibits a clear and consistent performance gain over baseline algorithms in almost all datasets, as indicated by the higher RI values. In particular, SFCM-PM often achieves the best or comparable performance to the best among the methods compared at each supervision level, showcasing its robustness and effectiveness.

Table 3. Comparison of RI values on standard datasets with 0% prior membership samples.

Dataset	k-Means	FCM	SFCM	cGFCM	SSFCM	SFCM-PM
Segment	0.822–	0.886=	0.886=	0.886=	0.886=	0.886
Iris	0.721–	0.832=	0.832=	0.832=	0.832=	0.832
Digits	0.707+	0.666=	0.667=	0.668=	0.667=	0.667
Water	0.587–	0.646–	0.675–	0.675–	0.675–	0.675
Air	0.705=	0.706=	0.706=	0.708=	0.706=	0.706
Banknote Authentication	0.532–	0.551=	0.551=	0.551=	0.551=	0.551
Parkinsons	0.518–	0.522=	0.522=	0.5–	0.522=	0.522
Ecoli	0.79=	0.794=	0.794=	0.797=	0.794=	0.794
+ / – / =	1/5/2	0/1/7	0/0/8	0/1/7	0/0/8	

Note: “+”, “–”, and “=” indicate that the results of the compared algorithms are significantly better than, worse than, and similar to that of SFCM-PM. In addition, the best metric values are highlighted in **bold**.

Table 4. Comparison of RI values on standard datasets with 5% prior membership samples.

Dataset	k-Means	FCM	SFCM	cGFCM	SSFCM	SFCM-PM
Segment	0.822–	0.886=	0.886=	0.885=	0.886=	0.887
Iris	0.721–	0.832=	0.832=	0.832=	0.832=	0.836
Digits	0.707–	0.666=	0.808–	0.818=	0.8–	0.824
Water	0.587–	0.645–	0.696–	0.708–	0.697–	0.725
Air	0.705–	0.706=	0.706=	0.709=	0.706=	0.711
Banknote Authentication	0.532–	0.551=	0.551=	0.537–	0.551=	0.559
Parkinsons	0.518–	0.522=	0.522=	0.5–	0.522=	0.541
Ecoli	0.79=	0.794=	0.794=	0.797=	0.794=	0.797
+ / – / =	0/7/1	0/3/5	0/3/5	0/3/5	0/3/5	

Note: “+”, “–”, and “=” indicate that the results of the compared algorithms are significantly better than, worse than, and similar to that of SFCM-PM. In addition, the best metric values are highlighted in **bold**.

Table 5. Comparison of RI values on standard datasets with 10% prior membership samples.

Dataset	k-Means	FCM	SFCM	cGFCM	SSFCM	SFCM-PM
Segment	0.822–	0.886=	0.886=	0.882=	0.886=	0.887
Iris	0.721–	0.832=	0.832=	0.832=	0.832=	0.835
Digits	0.707–	0.666–	0.839–	0.855=	0.831–	0.852
Water	0.587–	0.645–	0.716–	0.721–	0.719–	0.751
Air	0.705–	0.706–	0.706–	0.71=	0.706–	0.716
Banknote Authentication	0.532–	0.551–	0.551–	0.537–	0.551–	0.565
Parkinsons	0.518–	0.522–	0.523–	0.5–	0.524–	0.541
Ecoli	0.79=	0.794=	0.794=	0.797=	0.795=	0.8
+ / – / =	0/7/1	0/5/3	0/5/3	0/3/5	0/5/3	

Note: “+”, “–”, and “=” indicate that the results of the compared algorithms are significantly better than, worse than, and similar to that of SFCM-PM. In addition, the best metric values are highlighted in **bold**.

Table 6. Comparison of RI values on standard datasets with 15% prior membership samples.

Dataset	k-Means	FCM	SFCM	cGFCM	SSFCM	SFCM-PM
Segment	0.822–	0.886=	0.892=	0.886=	0.886=	0.892
Iris	0.721–	0.832=	0.832=	0.832=	0.832=	0.835
Digits	0.707–	0.666–	0.85=	0.864+	0.845–	0.856
Water	0.587–	0.645–	0.724–	0.737–	0.727–	0.758
Air	0.705–	0.706–	0.706–	0.709–	0.706–	0.717
Banknote Authentication	0.532–	0.551–	0.551–	0.537–	0.551–	0.568
Parkinsons	0.518–	0.522–	0.526–	0.501–	0.525–	0.542
Ecoli	0.79–	0.793–	0.795–	0.797–	0.795–	0.805
+ / – / =	0/8/0	0/6/2	0/5/3	1/5/2	0/6/2	

Note: “+”, “–”, and “=” indicate that the results of the compared algorithms are significantly better than, worse than, and similar to that of SFCM-PM. In addition, the best metric values are highlighted in **bold**.

Table 7. Comparison of RI values on standard datasets with 20% prior membership samples.

Dataset	k-Means	FCM	SFCM	cGFCM	SSFCM	SFCM-PM
Segment	0.822–	0.886–	0.894=	0.886–	0.886–	0.894
Iris	0.721–	0.832=	0.832=	0.832=	0.832=	0.833
Digits	0.707–	0.666–	0.867=	0.879+	0.858=	0.863
Water	0.587–	0.645–	0.745–	0.751–	0.746–	0.77
Air	0.705–	0.706–	0.706–	0.709–	0.706–	0.719
Banknote Authentication	0.532–	0.551–	0.551–	0.537–	0.551–	0.569
Parkinsons	0.518–	0.522–	0.529–	0.507–	0.529–	0.547
Ecoli	0.79–	0.793–	0.794–	0.797–	0.794–	0.808
+ / – / =	0/8/0	0/7/1	0/5/3	1/6/1	0/6/2	

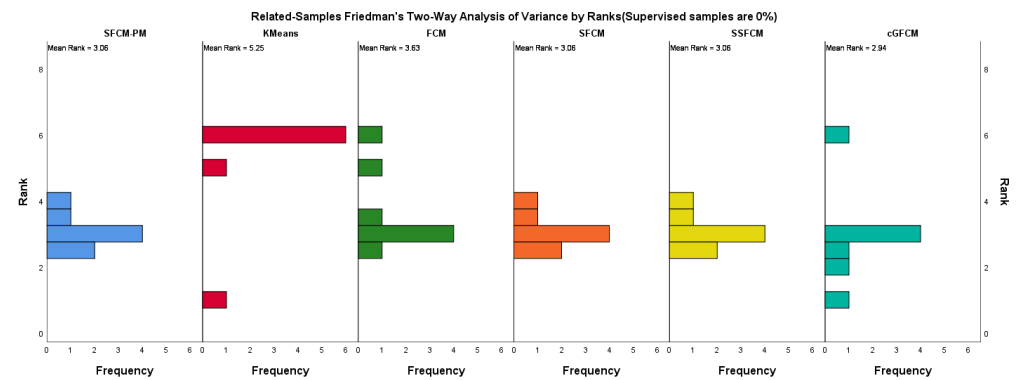
Note: “+”, “–”, and “=” indicate that the results of the compared algorithms are significantly better than, worse than, and similar to that of SFCM-PM. In addition, the best metric values are highlighted in **bold**.

4.4.2. Friedman Test

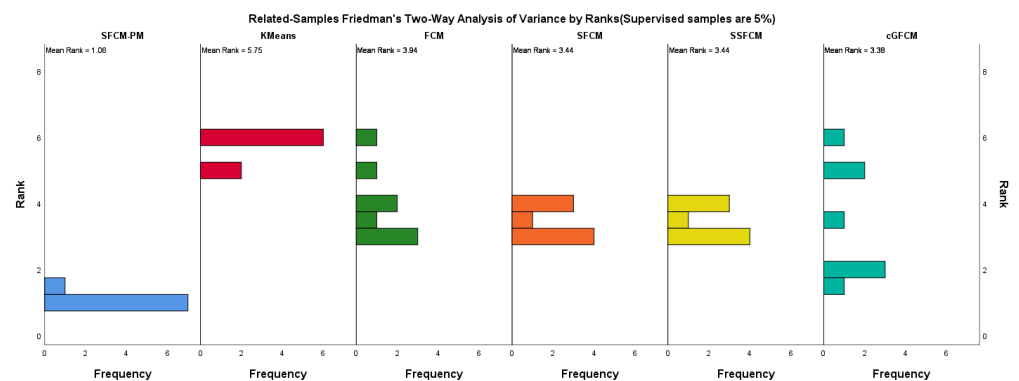
To further verify whether the SFCM-PM algorithm’s clustering performance on multiple datasets is significantly better than other comparison algorithms, a Friedman test was used for non-parametric statistical analysis. The Friedman test is a rank-sum test suitable for small samples and non-normally distributed data, and it is commonly used to compare whether there are significant differences in the average performance of multiple algorithms across multiple datasets. The results, visualized in Figure 2, offer a clear and compelling picture of SFCM-PM’s superiority.

The results presented in Figure 2 of the Friedman test are in line with the direct RI comparison results. At all regulatory levels (0% to 20%), the overall performance of SFCM-PM is superior to the other five algorithms. The average SFCM-PM ranking decreased significantly from 3.06 (at 0% prior samples) to approximately 1.13–1.31 (at 5–20% prior samples), clearly indicating that it can efficiently utilize the information from previous

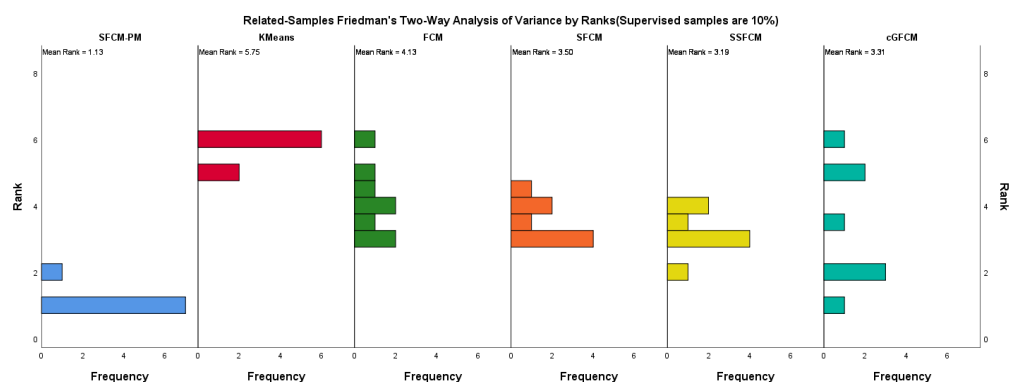
members and significantly improve the clustering results. In contrast, the k-means and FCM algorithms maintained relatively high and stable average rankings, which confirmed that they could not benefit from the provided regulation. The SFCM, SSFCM and cGFCM algorithms showed moderate improvements with increasing regulatory levels, but their average rankings were always higher than the SFCM-PM algorithm, further verifying the effectiveness of the prior information entropy criterion in the SFCM-PM algorithm.



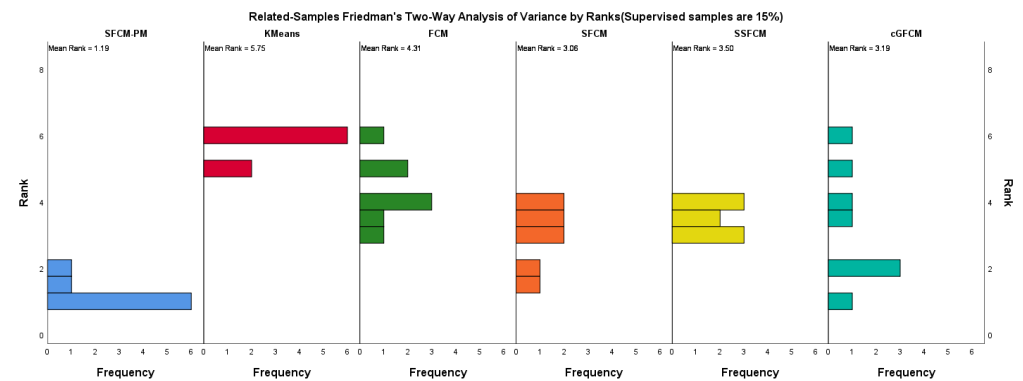
(a) Supervised samples are 0%



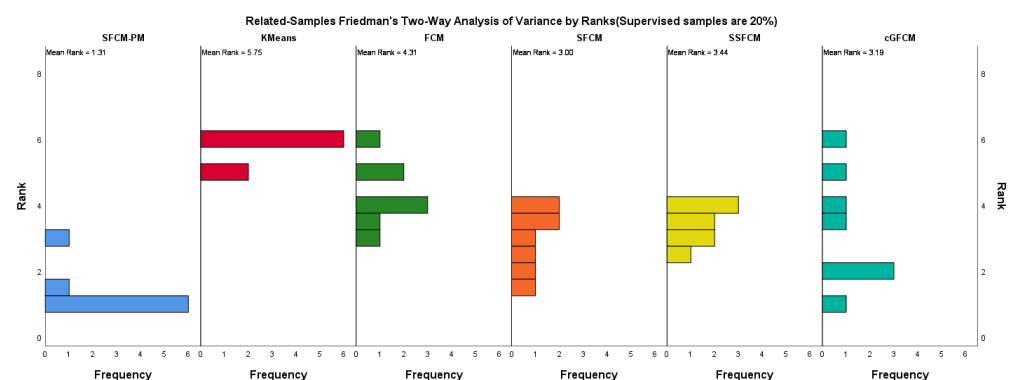
(b) Supervised samples are 5%



(c) Supervised samples are 10%



(d) Supervised samples are 15%



(e) Supervised samples are 20%

Figure 2. Friedman's two-way analysis of variance by rank results under different supervised sample proportions. (a) 0%, (b) 5%, (c) 10%, (d) 15%, and (e) 20%.

5. Discussion

For the eight datasets, 5%, 10%, 15%, and 20% of the samples were randomly selected in each dataset for prior membership degree processing. Tables 3–7 present the clustering performance of the SFCM-PM algorithm and other algorithms evaluated by RI metric under different proportions of prior samples. Furthermore, in order to verify the significance of the SFCM-PM algorithm well, we conducted the Friedman test during the experiment. The following is the attractive analysis results that we obtained:

- As shown in the tables, when the prior sample ratio is 0%, the SFCM-PM algorithm degenerates into an unsupervised fuzzy clustering method, with performance that is basically consistent with the FCM algorithm and shows no obvious advantages. After introducing prior information (prior membership sample ratios of 5% and above), the SFCM-PM algorithm significantly outperforms the other algorithms. In the Segment, Ecoli, Air, and Parkinsons datasets, when the prior sample ratio reaches 20%, the RI values are 0.894, 0.808, 0.719, and 0.547, respectively, significantly higher than other algorithms. This indicates that the auxiliary partitioning criterion based on prior information entropy proposed in the SFCM-PM algorithm can more effectively guide partition memberships to approach prior memberships, thereby improving clustering accuracy. On the Digits-389 dataset, the SFCM-PM algorithm performs relatively poorly, mainly due to the difficulty in distinguishing between classes caused by the visual similarity of digits. This prevents the algorithm from achieving ideal performance on this dataset, although it still outperforms traditional methods. After

introducing prior membership samples, the SFCM-PM algorithm's RI values on most datasets are superior to those of the other comparison algorithms, demonstrating its excellent clustering performance. As unsupervised methods, k-means and FCM exhibit stable but relatively low clustering accuracy throughout the experiment, with no significant changes due to the introduction of prior information. In contrast, semi-supervised fuzzy clustering methods such as SFCM, SSFCM, cGFCM, and SFCM-PM show varying degrees of performance improvement as the prior sample ratio increases, indicating that these algorithms can effectively model prior information and integrate it into the clustering process.

2. As illustrated in Figure 2, a systematic evaluation of the impact of varying proportions of supervised samples on the performance of the six clustering algorithms was conducted using the Friedman two-way ranked variance analysis. It is important to note that a lower average rank indicates superior algorithmic performance. The results demonstrate that SFCM-PM consistently exhibits significant and stable performance advantages across all levels of supervision, with its average rank maintaining a relatively low value. This indicates the algorithm's effective utilization of limited supervisory information to optimize clustering outcomes. When the proportion of prior samples is 0%, the average rank of SFCM-PM is 3.06, showing no marked advantage, which reflects the absence of prior information influence at this stage. As the proportion of prior samples increases, the performance gap between SFCM-PM and the other algorithms widens considerably, signifying its ability to efficiently integrate prior knowledge to enhance clustering performance. Specifically, at a 5% prior sample proportion, the average rank of SFCM-PM decreases to 1.06, significantly outperforming the other algorithms and demonstrating superior performance improvement. Moreover, only a small amount of supervision is sufficient to significantly boost algorithmic performance. In contrast, K-means and FCM algorithms, which are unable to effectively incorporate prior knowledge, maintain stable average ranks of 5.75 and 4.31 respectively, consistently ranking lower. Other algorithms such as SFCM, SSFCM, and cGFCM show minor fluctuations in average rank and exhibit moderate performance improvements as labeled data increase. Nevertheless, the average rank of SFCM-PM remains stable between 1.13 and 1.31, consistently occupying the top position, thereby validating its superiority and efficiency under limited prior supervision conditions.

6. Conclusions

In this paper, based on the classic FCM algorithm, a semi-supervised fuzzy clustering algorithm based on prior membership (SFCM-PM) is proposed by introducing prior information entropy to characterize the difference between partition membership and prior membership. The algorithm ensures that the similarity within the same cluster is maximized and the similarity between different clusters is minimized, while also making the partition memberships approach the prior memberships to the greatest extent possible. Numerical experiments on standard datasets and real-world datasets show that the proposed algorithm overcomes the limitation of traditional FCM algorithms that cannot utilize prior membership information and improves the precision of sample partitioning.

The SFCM-PM algorithm proposed in this study significantly outperforms various baseline clustering algorithms even when only a small number of samples (such as 5%) have prior membership information. By introducing prior information entropy as an auxiliary partitioning criterion, this algorithm demonstrates excellent robustness and consistency on both balanced and unbalanced datasets, as well as in practical applications. Its performance advantages under different supervision ratios have been verified by multiple statistical methods. Furthermore, the mathematical derivation and optimization process of SFCM-PM

is rigorous. When there is no prior information, the algorithm naturally degenerates into the traditional FCM, proving that it is a reasonable and effective extension. Although the SFCM-PM algorithm has achieved good results, there are still some limitations and areas for improvement. Future research can focus on comparing it with other advanced clustering algorithms, exploring the introduction of diverse types of prior knowledge (such as pairwise constraints and hierarchical information), and combining internal clustering validation indicators to achieve a more comprehensive assessment of clustering quality. At the same time, more in-depth empirical studies are needed to address the impact of noisy prior information on the algorithm's performance in order to enhance the algorithm's robustness and adaptability in actual complex environments.

Author Contributions: Conceptualization, Y.H. and G.Z.; data curation, G.M. and H.C.; formal analysis, H.Z.; funding acquisition, P.C. and Y.H.; investigation, P.C. and H.C.; methodology, Y.H. and G.Z.; project administration, P.C., J.Z. and H.C.; software, J.L.; supervision, Y.H., H.Z., P.C. and J.Z.; validation, J.L.; visualization, J.L. and G.M.; writing—original draft, G.Z., J.L. and G.M.; writing—review and editing, J.L., H.Z. and J.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the following funding sources: Key Research Project of Guangdong Province (2023B1111050010); Guangdong Provincial Soft Science Research Program Project (2025A1010010002); Guangdong Key Discipline Research Capacity Building Project (2024ZDJS055); Technology Planning Project of Guangdong Province, China (KTP20240831, KTP20240254); Science and Special Projects in Key Fields of Guangdong Provincial Universities (2023ZDZX3017, 2024ZDZX1085, and 2024ZDZX1031); 2022 Tertiary Bureau Education Scientific (202234607); Doctoral Starting Project of Provincial Department of Research Project of the Fund of Hanshan Normal Education Key Discipline Guangzhou University, Research Municipal China Capacity Education (QD202324); Enhancement Project Guang-(2021ZDJS058); and Guangdong Basic and Applied Basic Research Foundation (No. 2025A1515010466), project number: 2024A1010010001; project type: Guangdong Provincial Science and Technology Plan Project (Soft Science Research Field).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are available from the authors upon request.

Acknowledgments: During the preparation of this manuscript/study, the authors used ChatGPT (GPT-3.5) for the purposes of writing the graphical abstract. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

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

References

1. Abdelaziz, F.B.; Limam, O. Multiobjective fuzzy clustering with coalition formation: The case of brain image processing. *INFOR Inf. Syst. Oper. Res.* **2016**, *55*, 52–69. [\[CrossRef\]](#)
2. Anand, R.; Veni, S.; Aravinth, J. An application of image processing techniques for detection of diseases on brinjal leaves using k-means clustering method. In Proceedings of the 2016 International Conference on Recent Trends in Information Technology (ICRTIT), Chennai, India, 8–9 April 2016; pp. 1–6. [\[CrossRef\]](#)
3. Elmoataz, A.; Desquesnes, X.; Toutain, M. On the game p-Laplacian on weighted graphs with applications in image processing and data clustering. *Eur. J. Appl. Math.* **2017**, *28*, 1–27. [\[CrossRef\]](#)
4. Xu, Q.; Tang, D.; Cai, Q. Improved Fast Fuzzy C-Means Clustering Algorithm for Image Segmentation. *J. Nanjing Univ. Sci. Technol.* **2016**, *40*, 309–314.
5. Sun, Y.; Jiang, Z.; Shan, G.; Liu, H.; Rao, Y. Key Frame Extraction Based on Optimal Distance Clustering and Feature Fusion Expression. *J. Nanjing Univ. Sci. Technol.* **2018**, *42*, 416–423.
6. Qin, P.; Chen, W.; Zhang, M.; Li, D.; Feng, G. CC-GNN: A Clustering Contrastive Learning Network for Graph Semi-Supervised Learning. *IEEE Access* **2024**, *12*, 71956–71969. [\[CrossRef\]](#)

7. Frémal, S.; Lecron, F. Weighting Strategies for a Recommender System Using Item Clustering Based on Genres. *Expert Syst. Appl.* **2017**, *77*, 105–113. [\[CrossRef\]](#)
8. Forsati, R.; Doustdar, H. M.; Shamsfard, M.; Keikha, A.; Meybodi, M.R. A Fuzzy Co-Clustering Approach for Hybrid Recommender Systems. *Int. J. Hybrid Intell. Syst.* **2013**, *10*, 71–81. [\[CrossRef\]](#)
9. Selvi, C.; Sivasankar, E. A Novel Adaptive Genetic Neural Network (AGNN) Model for Recommender Systems Using Modified K-Means Clustering Approach. *Multimed. Tools Appl.* **2018**, *78*, 14303–14330. [\[CrossRef\]](#)
10. Anderberg, M.R. *Cluster Analysis for Applications*; Academic Press: New York, NY, USA, 1973; pp. 347–353.
11. Bezdek, J.C. Pattern Recognition with Fuzzy Objective Function Algorithms. *Adv. Appl. Pattern Recognit.* **1981**, *22*, 203–239.
12. Wagstaff, K.; Cardie, C.; Rogers, S.; Schrödl, S. Constrained K-means Clustering with Background Knowledge. In Proceedings of the Eighteenth International Conference on Machine Learning, Williamstown, MA, USA, 28 June–1 July 2001.
13. E. Yasunori, H. Yukihiro, Y. Makito and M. Sadaaki, On semi-supervised fuzzy c-means clustering. In Proceedings of the IEEE International Conference on Fuzzy Systems, Jeju, Republic of Korea, 20–24 August 2009; pp. 1119–1124. [\[CrossRef\]](#)
14. Yin, X.; Shu, T.; Huang, Q. Semi-supervised fuzzy clustering with metric learning and entropy regularization. *Knowl.-Based Syst.* **2012**, *35*, 304–311. [\[CrossRef\]](#)
15. Zhao, Q.; Zhan, S.; Cheng, R.; Zhu, J.; Zeng, H. A Benchmark for Vehicle Re-Identification in Mixed Visible and Infrared Domains. *IEEE Signal Process. Lett.* **2024**, *31*, 726–730. [\[CrossRef\]](#)
16. Zhao, K.; Dai, Y.; Jia, Z.; Ji, Y. General Fuzzy C-Means Clustering Strategy: Using Objective Function to Control Fuzziness of Clustering Results. *IEEE Trans. Fuzzy Syst.* **2022**, *30*, 3601–3616. [\[CrossRef\]](#)
17. Wang, Z.; Wang, S.-S.; Bai, L.; Wang, W.-S.; Shao, Y.-H. Semisupervised Fuzzy Clustering With Fuzzy Pairwise Constraints. *IEEE Trans. Fuzzy Syst.* **2022**, *30*, 3797–3811. [\[CrossRef\]](#)
18. Yao, L.; Yan, H.; Wei, Z. Robust Fuzzy Clustering Algorithm Based on Cross Entropy. *Appl. Res. Comput.* **2019**, *36*, 2948–2951.
19. Li, L.; Wang, F. Research on Semi-Supervised K-Medoids Algorithm Based on Hierarchical Strategy. *Appl. Res. Comput.* **2021**, *38*, 1387–1392.
20. Li, Q.; Yang, X.; Xie, X.; Liu, G. The Data Recovery Strategy on Machine Learning Against False Data Injection Attacks in Power Cyber Physical Systems. *Meas. Control* **2025**, *58*, 632–642. [\[CrossRef\]](#)
21. Tan, C.; Wu, H.; Tang, K.; Tan, C. An Extendable Gaussian Mixture Model for Lane-Based Queue Length Estimation Based on License Plate Recognition Data. *J. Adv. Transp* **2022**, *2022*, 5119209. [\[CrossRef\]](#)
22. Gao, J.; Tian, Y.-B.; Chen, X.-Z. Modeling of antenna resonant frequency based on co-training of semi-supervised Gaussian process with different kernel functions. *Int. J. RF Microw. Comput.-Aided Eng* **2021**, *31*, e22627. [\[CrossRef\]](#)
23. Chen, J.; Xu, J.; Gong, Y.; Xu, L. Ship Hull Principal Dimensions Optimization Employing Fuzzy Decision-Making Theory. *Math. Probl. Eng.* **2016**, *2016*, 5262160. [\[CrossRef\]](#)
24. Ren, Y.; Wang, C.; Li, B.; Yu, C.; Zhang, S. A Genetic Algorithm for Fuzzy Random and Low-Carbon Integrated Forward/Reverse Logistics Network Design. *Neural Comput. Applic.* **2020**, *32*, 2005–2025. [\[CrossRef\]](#)
25. Tang, H.; Wu, J.; Wu, F.; Chen, L.; Liu, Z.; Yan, H. An Optimization Framework for Collaborative Control of Power Loss and Voltage in Distribution Systems With DGs and EVs Using Stochastic Fuzzy Chance Constrained Programming. *IEEE Access* **2020**, *8*, 49013–49027. [\[CrossRef\]](#)
26. Feng, Y.; Liu, P.; Du, Y.; Jiang, Z. Cross Working Condition Bearing Fault Diagnosis Based on the Combination of Multimodal Network and Entropy Conditional Domain Adversarial Network. *J. Vib. Control* **2024**, *30*. [\[CrossRef\]](#)
27. Zhang, L.; Yang, T.; Xiao, G.; Xu, S. *MATLAB Data Analysis and Data Mining*; China Machine Press: Beijing, China, 2015.
28. Bai, F.J.; Gao, J.L.; Song, W.H.; He, S.Y. Research and improvement of semi-supervised fuzzy clustering algorithm. *Commun. Technol.* **2018**, *51*, 5.
29. Baghshah, M.S.; Shouraki, S.B. Kernel-Based Metric Learning for Semi-Supervised Clustering. *Neurocomputing* **2010**, *73*, 1352–1361. [\[CrossRef\]](#)
30. Grira, N. Active Semi-Supervised Fuzzy Clustering. *Pattern Recognit.* **2008**, *41*, 1834–1844. [\[CrossRef\]](#)
31. Huang, X.; Yang, X.; Zhao, J.; Xiong, L.; Ye, Y. A New Weighting k-Means Type Clustering Framework with an l2-Norm Regularization. *Knowl.-Based Syst.* **2018**, *151*, 165–179. [\[CrossRef\]](#)

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.