

Article

A Novel Semi-Supervised Fuzzy C-Means Clustering Algorithm Using Multiple Fuzzification Coefficients

Tran Dinh Khang ^{1,*} , Manh-Kien Tran ²  and Michael Fowler ² ¹ Department of Information Systems, Hanoi University of Science and Technology, Hanoi 10000, Vietnam² Department of Chemical Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada; kmtran@uwaterloo.ca (M.-K.T.); mfowler@uwaterloo.ca (M.F.)

* Correspondence: khangtd@soict.hust.edu.vn

Abstract: Clustering is an unsupervised machine learning method with many practical applications that has gathered extensive research interest. It is a technique of dividing data elements into clusters such that elements in the same cluster are similar. Clustering belongs to the group of unsupervised machine learning techniques, meaning that there is no information about the labels of the elements. However, when knowledge of data points is known in advance, it will be beneficial to use a semi-supervised algorithm. Within many clustering techniques available, fuzzy C-means clustering (FCM) is a common one. To make the FCM algorithm a semi-supervised method, it was proposed in the literature to use an auxiliary matrix to adjust the membership grade of the elements to force them into certain clusters during the computation. In this study, instead of using the auxiliary matrix, we proposed to use multiple fuzzification coefficients to implement the semi-supervision component. After deriving the proposed semi-supervised fuzzy C-means clustering algorithm with multiple fuzzification coefficients (sSMC-FCM), we demonstrated the convergence of the algorithm and validated the efficiency of the method through a numerical example.



Citation: Khang, T.D.; Tran, M.-K.; Fowler, M. A Novel Semi-Supervised Fuzzy C-Means Clustering Algorithm Using Multiple Fuzzification Coefficients. *Algorithms* **2021**, *14*, 258. <https://doi.org/10.3390/a14090258>

Academic Editor: Frank Werner

Received: 13 July 2021

Accepted: 27 August 2021

Published: 29 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 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: clustering technique; fuzzy C-means clustering; semi-supervised clustering; fuzzification coefficient; objective function

1. Introduction

Data clustering is a method that divides the elements of a data set into clusters such that data elements in the same cluster have similar properties, and data elements in different clusters have different properties [1,2]. Data clustering is an important pre-processing step to produce initial knowledge, supporting the decision-making for the next processing steps. In the current clustering methods, Fuzzy C-means clustering (FCM) algorithm gives relatively good results, taking advantage of the flexible nature of fuzzy logic [3]. In FCM, a data element can flexibly choose a cluster to belong to, characterized by the membership grade of the element in the cluster [4]. Specifically, the membership grade U_{ik} of element i belonging to cluster k has a value in the range from 0 to 1, and the larger the U_{ik} value, the more likely element i belongs to cluster k [5,6].

Clustering belongs to the class of unsupervised learning methods in which there is no given information about the labels of the data elements, dissimilar to classification methods. However, there are also cases in which knowledge of certain data can be known in advance; then, it becomes semi-supervised fuzzy clustering. For a semi-supervised clustering problem, there are often two goals: (i) clustering and labeling the data, and (ii) improving the quality of clustering based on the existing knowledge about the data [7–11]. The first goal is clustering, as defining cluster labels for the data remains the primary goal. With semi-supervision, the structure of the clusters and the cluster centers must still be clearly distinguished. The second goal is to improve the clustering quality based on the existing knowledge. The clustering quality can be evaluated from the clustering label. Of the two objectives, if more attention goes toward clustering labeling, then more supervisory

knowledge needs to be introduced, and the clustering results will be expected to be better. The disadvantage of this method is that certain measures of “similarity” may not be better, but this is acceptable when there are data points that are “shaped” by the knowledge; hence, the ability to update the cluster center is also improved [12]. There is knowledge about data that can be considered to be relatively accurate. It will be known to a degree in advance if a data element should or should not belong to a certain cluster. Such factors are dependent on the related expertise of the person conducting the clustering [13].

Yasunori et al. [14] proposed to improve the FCM algorithm by adding an auxiliary supervisory matrix, representing the supervised membership grade, into the objective function. Although the authors did not outline a specific method to determine the value of the supervisory matrix, the study made it clear that semi-supervised fuzzy clustering is completely feasible. The authors called this algorithm semi-supervised standard Fuzzy C-means clustering (sSFCM). This algorithm is mainly based on the approach that the supervised data point must belong to a certain cluster. The membership grade on that cluster then must be larger than the value in other unsupervised clusters.

The FCM method uses an exponential parameter m , also known as the fuzzification coefficient, in the objective function, which adjusts the membership grade U_{ik} of element i belonging to cluster k . In the standard FCM algorithm, the value of parameter m is selected from the beginning, for example, $m = 2$. In recent years, there have been studies that extended the selection of parameter m as an interval $[m_1, m_2]$ or a fuzzy value that, when reducing the type, would essentially be the selection of different $m \in [m_1, m_2]$ values, applied to each iteration [15]. The extension to use multiple fuzzification coefficients m instead of only one value in FCM was presented in [16]. When assigning each data point an appropriate fuzzification coefficient m based on the density of that point to the surrounding points, the clustering quality was improved.

This study is based on the concept of utilizing multiple fuzzification coefficients to apply to semi-supervised fuzzy clustering. When placing a supervised element i into cluster k , the appropriate fuzzification coefficient m_{ik} can be selected, which will be different from other fuzzification coefficients. This affects the calculation of U_{ik} which increases, and at the same time, $U_{ik} \hat{m}_{ik}$ changing also affects the determination of cluster k center. Similarly, when preventing a supervised element i from being in cluster k , its parameter m_{ik} will be different from other m_{ij} values. The adjustment of the fuzzification coefficient is similar to using the auxiliary matrix similar to other semi-supervised clustering methods, and it is also needed to ensure the convergence of the clustering algorithm. The contribution of this study is the proposal of a novel semi-supervised FCM algorithm using multiple fuzzification coefficients (sSMC-FCM) as well as the method to determine the appropriate fuzzification coefficient values for semi-supervision.

Section 2 of the paper outlines the FCM algorithm and the semi-supervised fuzzy clustering algorithm using auxiliary matrices. Section 3 presents the novel sSMC-FCM algorithm. Section 4 shows a numerical example using the proposed method and the subsequent results, while Section 5 presents concluding remarks.

2. Preliminaries

2.1. Standard Fuzzy C-Means Clustering (FCM) Algorithm

The standard FCM algorithm attempts to divide a finite number of N data elements $X = \{X_1, X_2, \dots, X_N\}$ into C clusters based on some given criteria. Each element $X_i \in X, i = 1, 2, \dots, N$, is a vector with D dimensions. The elements in X are divided into C clusters with cluster centers V_1, V_2, \dots, V_C in the centroid set V .

In FCM algorithm, U is a matrix that represents the membership of each element into each cluster. Matrix U has certain characteristics as below:

- U_{ik} is the membership grade of an element X_i in the cluster k with center V_k , where $1 \leq i \leq N; 1 \leq k \leq C$;
- $0 \leq U_{ik} \leq 1, 1 \leq i \leq N; 1 \leq k \leq C$ and $\sum_{j=1}^C U_{ij} = 1$, for each X_i ;
- The larger U_{ik} is, the more element X_i belongs in cluster k .

An objective function is defined such that the clustering algorithm must minimize the objective function (1):

$$J(U, V) = \sum_{i=1}^N \sum_{k=1}^C U_{ik}^m D_{ik}^2, \tag{1}$$

where $D_{ik}^2 = \|X_i - V_k\|^2$ is the distance between two vectors X_i and V_k , and m is the fuzzification coefficient of the algorithm.

Summary of steps for the standard FCM algorithm:

Input: the dataset $X = \{X_1, X_2, \dots, X_N\}$.

Output: the partition of X into C clusters.

- Step 1: Initialize value for V , let $l = 0$, set $\varepsilon > 0$ and $m > 1$.
- Step 2: At the l -th loop, update U according to the formula:

$$U_{ik} = \left(\sum_{j=1}^C \left(\frac{D_{ik}}{D_{ij}} \right)^{\frac{2}{m-1}} \right)^{-1} \tag{2}$$

- Step 3: Update V for the next step ($l + 1$), according to the formula:

$$V_k = \frac{\sum_{i=1}^N U_{ik}^m X_i}{\sum_{i=1}^N U_{ik}^m} \tag{3}$$

- Step 4: If $\|V^{(l)} - V^{(l+1)}\| < \varepsilon$, then go to Step 5; otherwise, let $l := l + 1$, and return to Step 2.
- Step 5: End.

2.2. Semi-Supervised Standard Fuzzy C-Means Clustering (sSFCM) Algorithms

The semi-supervised fuzzy clustering method in [14] added a supervisory matrix \bar{U} that represents the supervision of an element forced to belong or to not belong in a cluster. $\bar{U}_{ik} > 0$ when the supervised element i is placed into cluster k , and $\bar{U}_{ik} = 0$ when there is no supervision. This additional supervised membership grade \bar{U}_{ik} was added to the objective function to be minimized, and is shown as follows:

$$J(U, V) = \sum_{i=1}^N \sum_{k=1}^C |U_{ik} - \bar{U}_{ik}|^m D_{ik}^2, \tag{4}$$

where $\sum_{j=1}^C U_{ij} = 1$, $\sum_{j=1}^C \bar{U}_{ij} \leq 1$, $0 \leq U_{ik} \leq 1$, for all $1 \leq i \leq N$ and $1 \leq k \leq C$. The sSFCM algorithm, shown below, works to minimize $J(U, V)$ through many iterations.

Summary of steps for the sSFCM algorithm:

Input: the dataset $X = \{X_1, X_2, \dots, X_N\}$, and the supervised membership grade \bar{U} .

Output: the partition of X into C clusters.

- Step 1: Initialize value for V , let $l = 0$, set $\varepsilon > 0$ and $m > 1$.
- Step 2: At the l -th loop, update U according to the formula:

$$U_{ik} = \bar{U}_{ik} + \left(1 - \sum_{j=1}^C \bar{U}_{ij} \right) \left(\sum_{j=1}^C \left(\frac{D_{ik}}{D_{ij}} \right)^{\frac{2}{m-1}} \right)^{-1} \tag{5}$$

- Step 3: Update V for the next step ($l + 1$), according to the formula:

$$V_k = \frac{\sum_{i=1}^N |U_{ik} - \bar{U}_{ik}|^m X_i}{\sum_{i=1}^N |U_{ik} - \bar{U}_{ik}|^m} \tag{6}$$

- Step 4: If $\|V^{(l)} - V^{(l+1)}\| < \varepsilon$, then go to Step 5; otherwise, let $l := l + 1$, and return to Step 2.
- Step 5: End.

A numerical example of the sSFCM algorithm [14]:

A data set with 20 elements $X = \{X_1, X_2, \dots, X_{20}\}$, $X_i \in R^2$, as seen in Table 1, is to be divided into two clusters. Implement the sSFCM algorithm with $m = 2$ for 3 cases:

Table 1. Sample data set.

<i>i</i>	(<i>X</i> _{<i>i1</i>} , <i>X</i> _{<i>i2</i>})	<i>i</i>	(<i>X</i> _{<i>i1</i>} , <i>X</i> _{<i>i2</i>})	<i>i</i>	(<i>X</i> _{<i>i1</i>} , <i>X</i> _{<i>i2</i>})	<i>i</i>	(<i>X</i> _{<i>i1</i>} , <i>X</i> _{<i>i2</i>})
1	(0, 4.5)	6	(3.5, 5.5)	11	(9, 0)	16	(10, 0)
2	(0, 5.5)	7	(5.25, 4.5)	12	(9, 2.5)	17	(10, 2.5)
3	(1.75, 4.5)	8	(5.25, 5.5)	13	(9, 5)	18	(10, 5)
4	(1.75, 5.5)	9	(7, 4.5)	14	(9, 7.5)	19	(10, 7.5)
5	(3.5, 4.5)	10	(7, 5.5)	15	(9, 10)	20	(10, 10)

- Case 1: unsupervised, $\bar{U}_{ik} = 0$, for all $1 \leq i \leq 20, 1 \leq k \leq 2$ (standard FCM algorithm).
- Case 2: semi-supervised, attempting to place data points 9 and 10 into cluster 1, $\bar{U}_{9,1} = \bar{U}_{10,1} = 0.3$, otherwise, $\bar{U}_{ik} = 0$.
- Case 3: semi-supervised, attempting to place data points 9 and 10 into cluster 1, $\bar{U}_{9,1} = \bar{U}_{10,1} = 0.6$, otherwise, $\bar{U}_{ik} = 0$.

After clustering, we obtain matrices *U* for all 3 cases, as shown in Table 2.

Table 2. The resulting membership *U* for the 3 cases using the sSFCM algorithm.

<i>i</i>	Unsupervised	<i>i</i>	$\bar{U}_{9,1} = \bar{U}_{10,1} = 0.3$	<i>i</i>	$\bar{U}_{9,1} = \bar{U}_{10,1} = 0.6$
1	(0.91, 0.09)	1	(0.92, 0.08)	1	(0.92, 0.08)
2	(0.91, 0.09)	2	(0.92, 0.08)	2	(0.92, 0.08)
3	(0.98, 0.02)	3	(0.98, 0.02)	3	(0.98, 0.02)
4	(0.98, 0.02)	4	(0.98, 0.02)	4	(0.98, 0.02)
5	(0.97, 0.03)	5	(0.97, 0.03)	5	(0.97, 0.03)
6	(0.97, 0.03)	6	(0.97, 0.03)	6	(0.97, 0.03)
7	(0.68, 0.32)	7	(0.70, 0.30)	7	(0.72, 0.28)
8	(0.68, 0.32)	8	(0.70, 0.30)	8	(0.72, 0.28)
9	(0.19, 0.81)	9	(0.45, 0.55)	9	(0.69, 0.31)
10	(0.19, 0.81)	10	(0.45, 0.55)	10	(0.69, 0.31)
11	(0.28, 0.72)	11	(0.28, 0.72)	11	(0.28, 0.72)
12	(0.12, 0.88)	12	(0.12, 0.88)	12	(0.12, 0.88)
13	(0.00, 1.00)	13	(0.00, 1.00)	13	(0.00, 1.00)
14	(0.12, 0.88)	14	(0.12, 0.88)	14	(0.12, 0.88)
15	(0.28, 0.72)	15	(0.28, 0.72)	15	(0.28, 0.72)
16	(0.25, 0.75)	16	(0.25, 0.75)	16	(0.25, 0.75)
17	(0.11, 0.89)	17	(0.10, 0.90)	17	(0.10, 0.90)
18	(0.02, 0.98)	18	(0.01, 0.99)	18	(0.01, 0.99)
19	(0.11, 0.89)	19	(0.10, 0.90)	19	(0.10, 0.90)
20	(0.25, 0.75)	20	(0.25, 0.75)	20	(0.25, 0.75)

The clustering results indicate that sample points 9 and 10 belong to cluster 2 in case 1 and case 2 and to cluster 1 in case 3, showing the impact of the added semi-supervised component in the sSFCM algorithm through the supervised membership grade \bar{U} . Although in case 2, the attempt to place points 9 and 10 in cluster 1 was not successful, the points were much closer to cluster 1 compared to case 1.

In this study, instead of using the matrix \bar{U} , we propose the use of multiple fuzzification coefficients as the mean to make the original unsupervised FCM algorithm a semi-supervised one.

3. Semi-Supervised Fuzzy C-Means Clustering Algorithm with Multiple Fuzzification Coefficients (sSMC-FCM)

3.1. Derivation of the Proposed sSMC-FCM Algorithm

In this sub-section, we show the derivation of the proposed sSMC-FCM algorithm. Each membership U_{ik} can receive different fuzzification coefficients m_{ik} . The supervision of element i to belong to cluster k is performed through adjusting m_{ik} . A formula for m_{ik} can be given as follows, with exponential parameters $M, M' > 1$:

- $m_{ij} = M$ for all $1 \leq j \leq C$, for unsupervised elements i .
- $m_{ik} = M'$, and $m_{ij} = M$ for all $j \neq k$, for supervised elements i to belong to cluster k .

The approach to determine the exponential parameters M and M' is discussed in Section 3.2.

The objective function to be minimized in the sSMC-FCM algorithm is formulated as follows,

$$J(U, V) = \sum_{i=1}^N \sum_{k=1}^C U_{ik}^{m_{ik}} D_{ik}^2 \tag{7}$$

where $X = \{X_1, X_2, \dots, X_N\}$ is the dataset, N is the number of elements, C is the number of clusters, U_{ik} is the membership grade of an element X_i in the cluster k with center V_k , $0 \leq U_{ik} \leq 1, 1 \leq i \leq N, 1 \leq k \leq C, \sum_{j=1}^C U_{ij} = 1, \sum_{i=1}^N U_{ik} > 0, m_{ik} > 1$ is the fuzzification coefficient of X_i in cluster k , and $D_{ik}^2 = \|X_i - V_k\|^2$ is the distance between two vectors X_i and V_k .

Let $Y \subset \{1, 2, \dots, N\} \times \{1, 2, \dots, C\}$ be the set of supervised elements, we have:

$$m_{ij} = \begin{cases} M & , \quad \forall k : (i, k) \notin Y \\ M' & , \quad (i, j) \in Y \\ M & , \quad \exists k : (i, k) \in Y \wedge j \neq k \end{cases} \tag{8}$$

To solve the optimization problem shown in Equation (7), we utilize the Lagrange multiplier method. Let

$$L = \sum_{i=1}^N \sum_{k=1}^C U_{ik}^{m_{ik}} \|X_i - V_k\|^2 - \sum_{i=1}^N \lambda_i \left(\sum_{k=1}^C U_{ik} - 1 \right) \tag{9}$$

Moreover, with

$$\begin{cases} \frac{\partial L}{\partial U_{ik}} = 0, & 1 \leq i \leq N, 1 \leq k \leq C \\ \frac{\partial L}{\partial V_k} = 0, & 1 \leq k \leq C \end{cases}$$

we can obtain V_k using

$$\frac{\partial L}{\partial V_k} = \frac{\partial \left[\sum_{i=1}^N U_{ik}^{m_{ik}} (X_i - V_k)^2 \right]}{\partial V_k} = \sum_{i=1}^N U_{ik}^{m_{ik}} * 2 * (X_i - V_k) * (-1) = 0,$$

hence, we have

$$\sum_{i=1}^N U_{ik}^{m_{ik}} X_i = \sum_{i=1}^N U_{ik}^{m_{ik}} V_k,$$

or

$$V_k = \frac{\sum_{i=1}^N U_{ik}^{m_{ik}} X_i}{\sum_{i=1}^N U_{ik}^{m_{ik}}} \tag{10}$$

And next, we calculate U_{ik} using

$$\frac{\partial L}{\partial U_{ik}} = m_{ik} U_{ik}^{m_{ik}-1} D_{ik}^2 - \lambda_i = 0, \tag{11}$$

which implies the following:

$$U_{ik} = \left(\frac{\lambda_i}{m_{ik} D_{ik}^2} \right)^{\frac{1}{m_{ik}-1}} \tag{12}$$

Combining with $\sum_{j=1}^C U_{ij} = 1$, we consider the two following cases:

Case 1:

For unsupervised elements $i, m_{ij} = M$ for all $1 \leq j \leq C$, Equation (12) becomes:

$$U_{ik} = \left(\frac{\lambda_i}{MD_{ik}^2}\right)^{\frac{1}{M-1}} = \left(\frac{\lambda_i}{M}\right)^{\frac{1}{M-1}} \left(\frac{1}{D_{ik}^2}\right)^{\frac{1}{M-1}} = \left(\frac{\lambda_i}{M}\right)^{\frac{1}{M-1}} \left(\frac{1}{D_{ik}}\right)^{\frac{2}{M-1}} \tag{13}$$

Since $\sum_{j=1}^C U_{ij} = 1$, we have $\sum_{j=1}^C \left(\frac{\lambda_i}{MD_{ij}^2}\right)^{\frac{1}{M-1}} = 1$, or $\left(\frac{\lambda_i}{M}\right)^{\frac{1}{M-1}} \sum_{j=1}^C \left(\frac{1}{D_{ij}^2}\right)^{\frac{1}{M-1}} = 1$, or

$$\left(\frac{\lambda_i}{M}\right)^{\frac{1}{M-1}} = \frac{1}{\sum_{j=1}^C \left(\frac{1}{D_{ij}^2}\right)^{\frac{1}{M-1}}} = \frac{1}{\sum_{j=1}^C \left(\frac{1}{D_{ij}}\right)^{\frac{2}{M-1}}} \tag{14}$$

Replacing (14) into (13), we obtain

$$U_{ik} = \frac{1}{\sum_{j=1}^C \left(\frac{D_{ik}}{D_{ij}}\right)^{\frac{2}{M-1}}} = \left(\sum_{j=1}^C \left(\frac{D_{ik}}{D_{ij}}\right)^{\frac{2}{M-1}}\right)^{-1} \tag{15}$$

Case 2:

For supervised elements i to belong to cluster k , $m_{ik} = M'$ and $m_{ij} = M$ for $j \neq k$, Equation (11) becomes:

$$\begin{cases} M'U_{ik}^{M'-1}D_{ik}^2 = \lambda_i, & \text{and} \\ MU_{ij}^{M-1}D_{ij}^2 = \lambda_i, & j \neq k \end{cases}$$

or

$$M'U_{ik}^{M'-1}D_{ik}^2 = M'U_{i1}^{M'-1}D_{i1}^2 = \dots = M'U_{ij}^{M'-1}D_{ij}^2 = \dots = M'U_{iC}^{M'-1}D_{iC}^2$$

Combining with $\sum_{j=1}^C U_{ij} = 1$, to calculate U_{ij} , we need to solve the following:

$$\begin{cases} M'U_{ik}^{M'-1}D_{ik}^2 = MU_{i1}^{M-1}D_{i1}^2 = \dots = MU_{ij}^{M-1}D_{ij}^2 = \dots = MU_{iC}^{M-1}D_{iC}^2, & j \neq k \\ \sum_{j=1}^C U_{ij} = 1 & \text{for all } i \end{cases} \tag{16}$$

The steps to solve Equation (16) are shown through Equations (17)–(20) below. To make the presentation of the derivation seamless, we first give the calculation formulas, and the proof that this solution can solve Equation (16) will be presented in Section 3.2 with Proposition 1.

Specifically, we calculate $d_{min} = \min_{j=1, \dots, C} \{D_{ij}\}$, then:

$$d_{ij} = \frac{D_{ij}}{d_{min}}, \forall j = 1, \dots, C \tag{17}$$

$$\text{Calculate } \mu_{ij} = \left(\frac{1}{Md_{ij}^2}\right)^{\frac{1}{M-1}} \text{ for all } j \neq k \tag{18}$$

$$\text{Calculate } \mu_{ik} \text{ which is a variable in } \frac{\mu_{ik}}{\left(\mu_{ik} + \sum_{j=1, j \neq k}^C \mu_{ij}\right)^{\frac{M'-M}{M-1}}} = \left(\frac{1}{M'd_{ik}^2}\right)^{\frac{1}{M'-1}} \tag{19}$$

$$\text{Then normalize : } U_{ij} = \frac{\mu_{ij}}{\sum_{l=1}^C \mu_{il}} \tag{20}$$

From there, we have the sSMC-FCM algorithm as follows.

Summary of steps for the proposed sSMC-FCM algorithm:

Input: the dataset $X = \{X_1, \dots, X_N\}$, the fuzzification coefficients $M > 1$ and $M' > M$, and the set of supervised elements $Y \subset \{1, 2, \dots, N\} \times \{1, 2, \dots, C\}$.

Output: the partition of X into C clusters.

- Step 1: Initialize value for V , let $l = 0$, and set $\varepsilon > 0$.
- Step 2: At the l – *th* loop, update U according to Equation (15) for unsupervised elements, or according to Equations (17)–(20) for supervised elements.

- Step 3: Update V for the next step $(l + 1)$, according to Equation (10), with m_{ik} calculated using Equation (8).
- Step 4: If $\left\|V^{(l)} - V^{(l+1)}\right\| < \varepsilon$, then go to Step 5; otherwise, let $l := l + 1$, and return to Step 2.
- Step 5: End.

The sSMC-FCM algorithm has similar steps to the FCM and sSFCM algorithms, but it uses Equations (15) and (17)–(20) in Step 2 and Equations (8) and (10) in Step 3.

3.2. Determination of the Fuzzification Coefficients for Supervised Elements

In the proposed sSMC-FCM algorithm, when an element i is to be supervised, U_{ij} will be calculated according to Equations (17)–(20) in Step 2 of the algorithm. In this sub-section, we will discuss these formulas in more detail, as well as how to determine the exponential parameter M' for the supervised elements.

Proposition 1. U_{ij} calculated from Equations (17)–(20) satisfies Equation (16).

Proof of Proposition 1. From Equation (20), taking the sum of U_{ij} , we can see that it satisfies the condition of $\sum_{j=1}^C U_{ij} = 1$ in Equation (16). \square

Next, we calculate $M'U_{ik}^{M'-1}D_{ik}^2$ by gradually replacing from the formulas (20), (19), (18), (17), as follows:

$$\begin{aligned}
 M'U_{ik}^{M'-1}D_{ik}^2 &= M' \left(\frac{\mu_{ik}}{\sum_{l=1}^C \mu_{il}} \right)^{M'-1} (d_{ik}d_{min})^2 \\
 &= M' \left(\frac{\left(\frac{1}{M'd_{ik}^2} \right)^{\frac{1}{M'-1}} (\mu_{ik} + \sum_{j=1, j \neq k}^C \mu_{ij})^{\frac{M'-M}{M'-1}}}{\sum_{l=1}^C \mu_{il}} \right)^{M'-1} (d_{ik}d_{min})^2 \\
 &= M' \left(\frac{\left(\frac{1}{M'd_{ik}^2} \right)^{\frac{1}{M'-1}} (\sum_{j=1}^C \mu_{ij})^{\frac{M'-M}{M'-1}}}{\sum_{l=1}^C \mu_{il}} \right)^{M'-1} d_{ik}^2 d_{min}^2 \\
 &= M' \frac{\frac{1}{M'd_{ik}^2} (\sum_{j=1}^C \mu_{ij})^{M'-M}}{(\sum_{l=1}^C \mu_{il})^{M'-1}} d_{ik}^2 d_{min}^2 = \frac{d_{min}^2}{(\sum_{l=1}^C \mu_{il})^{M-1}}
 \end{aligned}
 \tag{21}$$

Similarly, calculate $MU_{ij}^{M-1}D_{ij}^2$ for all $j \neq k$

$$\begin{aligned}
 MU_{ij}^{M-1}D_{ij}^2 &= M \left(\frac{\mu_{ij}}{\sum_{l=1}^C \mu_{il}} \right)^{M-1} (d_{ij}d_{min})^2 \\
 &= M \left(\frac{\left(\frac{1}{Md_{ij}^2} \right)^{\frac{1}{M-1}}}{\sum_{l=1}^C \mu_{il}} \right)^{M-1} d_{ij}^2 d_{min}^2 = \frac{d_{min}^2}{(\sum_{l=1}^C \mu_{il})^{M-1}}
 \end{aligned}
 \tag{22}$$

From (21) and (22), it can be seen that $M'U_{ik}^{M'-1}D_{ik}^2 = MU_{i1}^{M-1}D_{i1}^2 = \dots = MU_{ij}^{M-1}D_{ij}^2 = \dots = MU_{ic}^{M-1}D_{ic}^2$ for all $j \neq k$. Therefore, combining with $\sum_{j=1}^C U_{ij} = 1$, Equation (16) is satisfied, and Proposition 1 is proven.

We can hence apply Equations (17)–(20) in Step 2 of the proposed algorithm.

Next, we investigate how to determine the exponential parameter M' . We will utilize Proposition 2 after proving it.

Proposition 2. The function $f(x) = \left(\frac{1}{ax}\right)^{\frac{1}{x-1}}$ with $a \geq 1$ is an increasing function when $x > 1$.

Proof of Proposition 2. Consider the function $f(x) = \left(\frac{1}{ax}\right)^{\frac{1}{x-1}} = e^{\ln\left(\left(\frac{1}{ax}\right)^{\frac{1}{x-1}}\right)} = e^{\frac{1}{x-1}\ln\left(\frac{1}{ax}\right)}$, Calculate its derivative:

$$\begin{aligned} f'(x) &= e^{\frac{1}{x-1}\ln\left(\frac{1}{ax}\right)} \cdot \left(\left(\frac{1}{x-1}\right)' \ln\left(\frac{1}{ax}\right) + \left(\frac{1}{x-1}\right) \ln\left(\frac{1}{ax}\right)' \right) \\ &= e^{\frac{1}{x-1}\ln\left(\frac{1}{ax}\right)} \cdot \left(\frac{-1}{(x-1)^2}(-1)\ln(ax) + \left(\frac{1}{x-1}\right) \frac{1}{ax} \left(\frac{1}{ax}\right)' \right) \\ &= e^{\frac{1}{x-1}\ln\left(\frac{1}{ax}\right)} \cdot \left(\frac{\ln(ax)}{(x-1)^2} - \frac{1}{x(x-1)} \right) = e^{\frac{1}{x-1}\ln\left(\frac{1}{ax}\right)} \cdot \frac{x\ln(ax) - x + 1}{x(x-1)^2} \end{aligned}$$

For $x > 1, a \geq 1$, in $f'(x)$, we have $e^{\frac{1}{x-1}\ln\left(\frac{1}{ax}\right)} > 0, x\ln(ax) - x + 1 \geq x\ln x - x + 1$, and $x(x-1)^2 > 0$. For $g(x) = x\ln x - x + 1$, we have $g'(x) = \ln x + \frac{x}{x} - 1 = \ln x > 0$ for all $x > 1$. Therefore, $g(x)$ is an increasing function in the range $[1, +\infty)$, and thus, $g(x) > g(1) = 0$, and hence, $x\ln(ax) - x + 1 \geq x\ln x - x + 1 > 0$. Overall, all the components of $f'(x)$ are greater than 0. Therefore, $f'(x) > 0$, or $f(x)$ is an increasing function, for all $x > 1$. □

Applying Proposition 2 to the following formula used to calculate M' for the supervised elements, $f(M') = \left(\frac{1}{M'd_{ik}^2}\right)^{\frac{1}{M'-1}}$, where $d_{ik}^2 \geq 1, M' > 1$, we see that this function is an increasing function. Therefore, as M' increases, the membership grade U_{ik} increases. Setting the supervision for an element corresponds to selecting a larger fuzzification coefficient M' for that element compared to other fuzzification coefficients.

Then, Proposition 3 is utilized to guide the determination of parameter M' .

Proposition 3. Let U_{ik} be the membership grade of element X_i in its semi-supervised placement into cluster k according to Equations (17)–(20) with a given parameter M and a to-be-determined parameter M' . Let U'_{ik} be the membership grade of element X_i in its unsupervised placement into cluster k according to Equation (15). With $\alpha \in (0, 1)$, we have $U_{ik} \geq \alpha$ if the following equation is satisfied:

$$M'\alpha^{M'-1} \leq M \left(\frac{1-\alpha}{\frac{1}{U'_{ik}} - 1} \right)^{M-1} \tag{23}$$

Proof of Proposition 3. From Equation (15), we have

$$U'_{ik} = \frac{1}{\sum_{j=1}^C \left(\frac{D_{ik}}{D_{ij}}\right)^{\frac{2}{M-1}}} = \frac{1}{\left(\frac{D_{ik}}{D_{ik}}\right)^{\frac{2}{M-1}} + \sum_{j=1, j \neq k}^C \left(\frac{D_{ik}}{D_{ij}}\right)^{\frac{2}{M-1}}} = \frac{1}{1 + \sum_{j=1, j \neq k}^C \left(\frac{D_{ik}}{D_{ij}}\right)^{\frac{2}{M-1}}}$$

From Equation (17), we have $\frac{d_{ik}}{d_{ij}} = \frac{D_{ik}}{D_{ij}}$, hence, $U'_{ik} = \frac{1}{1 + \sum_{j=1, j \neq k}^C \left(\frac{d_{ik}}{d_{ij}}\right)^{\frac{2}{M-1}}}$.

Substitute into (23), we have the equivalent

$$M'\alpha^{M'-1} \leq \frac{M(1-\alpha)^{M-1}}{\left(\sum_{j=1, j \neq k}^C \left(\frac{d_{ik}}{d_{ij}}\right)^{\frac{2}{M-1}}\right)^{M-1}} = \frac{(1-\alpha)^{M-1}}{d_{ik}^2 \left(\sum_{j=1, j \neq k}^C \left(\frac{1}{Md_{ij}^2}\right)^{\frac{1}{M-1}}\right)^{M-1}}$$

or,

$$\frac{1}{M'd_{ik}^2} \geq \frac{\alpha^{M'-1}}{(1-\alpha)^{M-1}} \left(\sum_{j=1, j \neq k}^C \left(\frac{1}{Md_{ij}^2}\right)^{\frac{1}{M-1}} \right)^{M-1}$$

Substitute in μ_{ij} from Equation (18), we have $\frac{1}{M'd_{ik}^2} \geq \frac{\alpha^{M'-1}}{(1-\alpha)^{M-1}} \left(\sum_{j=1, j \neq k}^C \mu_{ij}\right)^{M-1}$.

Substitute $\frac{1}{M'd_{ik}^2}$ in the above with Equation (19), we have

$$\frac{\mu_{ik}^{M'-1}}{\left(\mu_{ik} + \sum_{j=1, j \neq k}^C \mu_{ij}\right)^{M'-M}} \geq \frac{\alpha^{M'-1}}{(1-\alpha)^{M-1}} \left(\sum_{j=1, j \neq k}^C \mu_{ij}\right)^{M-1} \tag{24}$$

Next, from Equation (20), we have $U_{ik} = \frac{\mu_{ik}}{\sum_{j=1}^C \mu_{ij}} = \frac{\mu_{ik}}{\mu_{ik} + \sum_{j=1, j \neq k}^C \mu_{ij}}$, hence $\sum_{j=1, j \neq k}^C \mu_{ij} = \frac{\mu_{ik}}{U_{ik}} - \mu_{ik}$, which is substituted into Equation (24) to have

$$\frac{\mu_{ik}^{M'-1}}{\left(\frac{\mu_{ik}}{U_{ik}}\right)^{M'-M}} \geq \frac{\alpha^{M'-1}}{(1-\alpha)^{M-1}} \left(\frac{\mu_{ik}(1-U_{ik})}{U_{ik}}\right)^{M-1}$$

Since $M' > M$, and from the above, we have $\mu_{ik}^{M'-1} U_{ik}^{M'-M} \geq \frac{\alpha^{M'-1} \mu_{ik}^{M-1}}{(1-\alpha)^{M-1}} \left(\frac{1-U_{ik}}{U_{ik}}\right)^{M-1}$.

Since μ_{ik} is to be solved in Equation (19), $\mu_{ik} > 0$, and the above can be simplified into $\frac{U_{ik}^{M'-M} U_{ik}^{M-1}}{(1-U_{ik})^{M-1}} \geq \frac{\alpha^{M'-1}}{(1-\alpha)^{M-1}}$ or

$$\frac{U_{ik}^{M'-1}}{(1-U_{ik})^{M-1}} \geq \frac{\alpha^{M'-1}}{(1-\alpha)^{M-1}} \tag{25}$$

Therefore, if Equation (23) is satisfied, then Equation (25) is also satisfied.

Consider the function $f(x) = \frac{x^{M'-1}}{(1-x)^{M-1}}$ with $x \in (0, 1)$ we have its derivative

$$\begin{aligned} f'(x) &= \frac{(M'-1)x^{M'-2}(1-x)^{M-1} - x^{M'-1}(M-1)(1-x)^{M-2}(-1)}{(1-x)^{2(M-1)}} \\ &= \frac{(1-x)^{M-2}x^{M'-2}((M'-1)(1-x) + x(M-1))}{(1-x)^{2(M-1)}} \\ &= \frac{x^{M'-2}(M'-Mx+Mx-1)}{(1-x)^M} \end{aligned}$$

Since $M' > M$, and from the above derivative, with $x \in (0, 1)$, we have

$$M' - M'x + Mx - 1 \geq M' - M'x + Mx - M = M'(1-x) - M(1-x) = (M' - M)(1-x) > 0$$

Since $x^{M'-2} > 0$ and $(1-x)^M > 0$, hence, $f'(x) > 0$, and therefore, $f(x)$ is an increasing function. Equation (25) has the form of $f(U_{ik}) \geq f(\alpha)$, so then $U_{ik} \geq \alpha$. □

From Proposition 3, Equation (23) demonstrates an approach to determine parameter M' . The right-hand side of (23) is a value that can be calculated knowing M , α , and the distances, while the left-hand side of (23) is a decreasing function as M' increases. From this, we can start at $M' = M$, and then increase M' value and check simultaneously until Equation (23) is satisfied to solve for M' .

Example: Given the data set from Table 1, the element X_9 has $U_{91} = 0.189$, with $M = 2$. If we want to apply supervision such that $U_{91} \geq 0.5$ to put X_9 into cluster 1, then from Equation (20), we need to determine M' such that the condition $0.5^{M'-1} M' \leq 0.233$ is satisfied. For $M' = 2$, $0.5^{M'-1} M' = 1$, and hence it does not satisfy the condition. We will have to increase the value of M' to decrease the value of $0.5^{M'-1} M'$ to satisfy the condition. Through certain iterations, we obtain $M' = 5.582$, and $0.5^{M'-1} M' = 0.233$, which satisfies the condition, and effectively places X_9 into cluster 1.

A point that can be further discussed is that in Step 2 of the sSMC-FCM algorithm, it is necessary to solve Equation (19) to obtain μ_{ik} . In Equation (19), its right-hand side has a value in the range $[0,1]$. Set $A = \sum_{j=1, j \neq k}^C \mu_{ij}$, $b = \frac{M'-M}{M'-1}$, with $b < 1$, the left-hand side of this equation has the form of $f(x) = \frac{x}{(x+A)^b}$, which has the derivative of $f'(x) = \frac{(x+A)^b - xb(x+A)^{b-1}}{(x+A)^{2b}} = \frac{x+A-xb}{(x+A)^{b+1}} = \frac{(1-b)x+A}{(x+A)^{b+1}} > 0$ for $x > 0$. From this, we can see that the left-hand side of Equation (19) is an increasing function, while the right-hand side has a value in the range $[0,1]$. Therefore, we can use an approximation method to obtain μ_{ik} , starting from $\mu_{ik} = 0$, and then increasing gradually to determine the solution.

4. Numerical Examples

To evaluate the proposed algorithm, we use the same numerical example shown in Section 2.2. A data set with 20 elements $X = \{X_1, X_2, \dots, X_{20}\}$, $X_i \in R^2$, as seen in Table 1, is to be divided into two clusters. Implement the sSMC-FCM algorithms with $M = 2$ for the following 3 cases:

- Case 1: unsupervised, $M' = M = 2$ (standard FCM algorithm).
- Case 2: semi-supervised, attempting to place data points 9 and 10 into cluster 1, $M' = 4$, $M = 2$.
- Case 3: semi-supervised, attempting to place data points 9 and 10 into cluster 1, $M' = 8$, $M = 2$.

The Euclidian distance was used in the calculations and results in this work. Table 3 shows membership matrices U for each case, as the results of clustering.

Table 3. The resulting membership U for the 3 cases using the sSMC-FCM algorithm.

i	Unsupervised	i	$M'=4$	i	$M'=8$
1	(0.91, 0.09)	1	(0.92, 0.08)	1	(0.92, 0.08)
2	(0.91, 0.09)	2	(0.92, 0.08)	2	(0.92, 0.08)
3	(0.98, 0.02)	3	(0.98, 0.02)	3	(0.98, 0.02)
4	(0.98, 0.02)	4	(0.98, 0.02)	4	(0.98, 0.02)
5	(0.97, 0.03)	5	(0.98, 0.02)	5	(0.98, 0.02)
6	(0.97, 0.03)	6	(0.98, 0.02)	6	(0.98, 0.02)
7	(0.68, 0.32)	7	(0.71, 0.29)	7	(0.71, 0.29)
8	(0.68, 0.32)	8	(0.71, 0.29)	8	(0.71, 0.29)
9	(0.19, 0.81)	9	(0.43, 0.57)	9	(0.60, 0.40)
10	(0.19, 0.81)	10	(0.43, 0.57)	10	(0.60, 0.40)
11	(0.28, 0.72)	11	(0.28, 0.72)	11	(0.28, 0.72)
12	(0.12, 0.88)	12	(0.12, 0.88)	12	(0.12, 0.88)
13	(0.00, 1.00)	13	(0.00, 1.00)	13	(0.00, 1.00)
14	(0.12, 0.88)	14	(0.12, 0.88)	14	(0.12, 0.88)
15	(0.28, 0.72)	15	(0.28, 0.72)	15	(0.28, 0.72)
16	(0.25, 0.75)	16	(0.25, 0.75)	16	(0.25, 0.75)
17	(0.11, 0.89)	17	(0.11, 0.89)	17	(0.10, 0.90)
18	(0.02, 0.98)	18	(0.01, 0.99)	18	(0.01, 0.99)
19	(0.11, 0.89)	19	(0.11, 0.89)	19	(0.10, 0.90)
20	(0.25, 0.75)	20	(0.25, 0.75)	20	(0.25, 0.75)

From the results in Table 3, we have the following observations:

- As M' increases, the membership grades of the supervised elements increase. Initially, with no supervision, data points 9 and 10 were placed into cluster 2. With supervision and $M' = 4$, their membership grades increased but not enough to move into cluster 1, while with $M' = 8$, these two data points were successfully placed into cluster 1;
- In general, the sSMC-FCM algorithm converges after a similar number of iterations to the standard FCM algorithm. For instance, with $\epsilon = 10^{-3}$, it required about 10 iterations;
- When M' is changed, the cluster centers change to suit the supervised elements. For instance, case 1 has cluster centers $V_1 = (2.719, 5.000)$, $V_2 = (9.018, 5.000)$, case 2 has cluster centers $V_1 = (2.735, 5.000)$, $V_2 = (9.200, 5.000)$, and case 3 has cluster centers

$V_1 = (2.714, 5.000)$, $V_2 = (9.303, 5.000)$. We can observe that, from case 1 to 2, the cluster center V_2 moves further away from data points 9 and 10;

- Compared to the sSFCM algorithm, we can see that the matrices U in the case of $M' = 4$, $M = 2$ using the sSMC-FCM algorithm shown in Table 3 are similar to the corresponding matrices U in the case of $\bar{U}_{9,1} = \bar{U}_{10,1} = 0.3$ using the sSFCM algorithm shown in Table 2. Both cases were able to increase the membership grade of the points to belong to cluster 1 but were not successfully in moving the points into cluster 1;
- In future works, it is possible to expand on other different representations for the fuzzification coefficients, combining with hedge algebra [17–19] in new representations and calculations.

5. Conclusions

In this study, we developed a novel clustering algorithm, called sSMC-FCM, based on the standard FCM algorithm, adding the semi-supervision aspect through the use of multiple fuzzification coefficients or also known as exponential parameters. In the sSMC-FCM algorithm, we allow the data elements to have different fuzzification coefficients, instead of only one such as in the standard FCM method. The expansion from hard clustering to fuzzy clustering involves the addition of fuzzification coefficients; hence, the determination of the fuzzification coefficient values is an interesting topic to be further researched. In this study, we derived the novel sSMC-FCM algorithm and proved three propositions to show the convergence of the algorithm and to explain how to determine the fuzzification coefficients and demonstrated the efficiency of the algorithm using a numerical example. The proposed algorithm added supervision in the normally unsupervised FCM clustering algorithm. This method can be applied in practical problems such as remote sensing image segmentation, when knowing an image is certainly of an image type such as a lake, but due to image noises and the unsupervised approach, the attribute features may cause that image to be perceived as another image type, such as clouds, leading to undesirable results. With the proposed semi-supervised method, we should be able to place the image into the knowingly correct image type as desired.

Author Contributions: Conceptualization, T.D.K.; methodology, T.D.K. and M.-K.T.; software, T.D.K.; validation, T.D.K., M.-K.T., and M.F.; formal analysis, T.D.K.; writing—original draft preparation, T.D.K. and M.-K.T.; writing—review and editing, M.F.; supervision, T.D.K. and M.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Vietnam National Foundation for Science and Technology Development (NAFOSTED) under Grant Number 102.05–2018.02.

Institutional Review Board Statement: Not application.

Informed Consent Statement: Not application.

Data Availability Statement: Not application.

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

References

1. Arora, J.; Khatter, K.; Tushir, M. Fuzzy c-Means Clustering Strategies: A Review of Distance Measures. *Softw. Eng.* **2018**, 153–162.
2. Everitt, B.S.; Landau, S.; Leese, M.; Stahl, D. *Cluster Analysis*, 5th ed.; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2011.
3. Havens, T.C.; Bezdek, J.C.; Leckie, C.; Hall, L.O.; Palaniswami, M. Fuzzy c-Means Algorithms for Very Large Data. *IEEE Trans. Fuzzy Syst.* **2012**, *20*, 1130–1146. [[CrossRef](#)]
4. Gosain, A.; Dahiya, S. Performance Analysis of Various Fuzzy Clustering Algorithms: A Review. *Procedia Comput. Sci.* **2016**, *79*, 100–111. [[CrossRef](#)]
5. Ruspini, E.H.; Bezdek, J.C.; Keller, J.M. Fuzzy Clustering: A Historical Perspective. *IEEE Comput. Intell. Mag.* **2019**, *14*, 45–55. [[CrossRef](#)]
6. Vendramin, L.; Campello, R.J.G.B.; Hruschka, E.R. Relative Clustering Validity Criteria: A Comparative Overview. *Stat. Anal. Data Min.* **2010**, *3*, 209–235. [[CrossRef](#)]
7. Casalino, G.; Castellano, G.; Mencar, C. Data stream classification by dynamic incremental semi-supervised fuzzy clustering. *Int. J. Artif. Intell. Tools* **2019**, *28*, 1960009. [[CrossRef](#)]

8. Gan, H.; Fan, Y.; Luo, Z.; Huang, R.; Yang, Z. Confidence-weighted safe semi-supervised clustering. *Eng. Appl. Artif. Intell.* **2019**, *81*, 107–116. [[CrossRef](#)]
9. Mai, S.D.; Ngo, L.T. Multiple kernel approach to semi-supervised fuzzy clustering algorithm for land-cover classification. *Eng. Appl. Artif. Intell.* **2018**, *68*, 205–213. [[CrossRef](#)]
10. Komori, O.; Eguchi, S. A Unified Formulation of k-Means, Fuzzy c-Means and Gaussian Mixture Model by the Kolmogorov–Nagumo Average. *Entropy* **2021**, *23*, 518. [[CrossRef](#)] [[PubMed](#)]
11. Son, L.H.; Tuan, T.M. Dental segmentation from X-ray images using semi-supervised fuzzy clustering with spatial constraints. *Eng. Appl. Artif. Intell.* **2017**, *59*, 186–195. [[CrossRef](#)]
12. Maraziotis, I.A. A semi-supervised fuzzy clustering algorithm applied to gene expression data. *Pattern Recognit.* **2012**, *45*, 637–648. [[CrossRef](#)]
13. Śmieja, M.; Struski, Ł.; Figueiredo, M.A. A classification-based approach to semi-supervised clustering with pairwise constraints. *Neural Netw.* **2020**, *127*, 193–203. [[CrossRef](#)] [[PubMed](#)]
14. Yasunori, E.; Yukihiro, H.; Makito, Y.; Yasunori, M.S. On semi-supervised fuzzy c-means clustering. In Proceedings of the IEEE International Conference on Fuzzy Systems, Jeju Island, Korea, 20–24 August 2009.
15. Hwang, C.; Rhee, F.C.-H. Uncertain Fuzzy Clustering: Interval Type-2 Fuzzy Approach to C-Means. *IEEE Trans. Fuzzy Syst.* **2007**, *15*, 107–120. [[CrossRef](#)]
16. Khang, T.D.; Vuong, N.D.; Tran, M.-K.; Fowler, M. Fuzzy C-Means Clustering Algorithm with Multiple Fuzzification Coefficients. *Algorithms* **2020**, *13*, 158. [[CrossRef](#)]
17. Khang, T.D.; Phong, P.A.; Dong, D.K.; Trang, C.M. Hedge Algebraic Type-2 Fuzzy Sets. In Proceedings of the Conference: FUZZ-IEEE 2010, IEEE International Conference on Fuzzy Systems, Barcelona, Spain, 18–23 July 2010.
18. Nguyen, C.H.; Tran, D.K.; Nam, H.V.; Nguyen, H.C. Hedge Algebras, Linguistic-Valued Logic and Their Application to Fuzzy Reasoning. *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.* **1999**, *7*, 347–361. [[CrossRef](#)]
19. Phong, P.A.; Khang, T.D.; Dong, D.K. A fuzzy rule-based classification system using Hedge Algebraic Type-2 Fuzzy Sets. In Proceedings of the Annual Conference of the North American Fuzzy Information Processing Society (NAFIPS), El Paso, TX, USA, 31 October–4 November 2016; pp. 265–270.