



Article

New Classifier Ensemble and Fuzzy Community Detection Methods Using POP Choquet-like Integrals

Xiaohong Zhang^{1,2}, Haojie Jiang^{1,*} and Jingqian Wang¹

¹ School of Mathematics and Data Science, Shaanxi University of Science and Technology, Xi'an 710016, China; zhangxiaohong@sust.edu.cn (X.Z.); wangjingqianw@163.com (J.W.)

² Shaanxi Joint Laboratory of Artificial Intelligence, Shaanxi University of Science and Technology, Xi'an 710016, China

* Correspondence: jiangjiang202101@163.com

Abstract: Among various data analysis methods, classifier ensemble (data classification) and community network detection (data clustering) have aroused the interest of many scholars. The maximum operator, as the fusion function, was always used to fuse the results of the base algorithms in the classifier ensemble and the membership degree of nodes to classes in the fuzzy community. It is vital to use generalized fusion functions in ensemble and community applications. Since the Pseudo overlap function and the Choquet-like integrals are two new fusion functions, they can be combined as a more generalized fusion function. Along this line, this paper presents new classifier ensemble and fuzzy community detection methods using a pseudo overlap pair (POP) Choquet-like integral (expressed as a fraction). First, the pseudo overlap function pair is proposed to replace the product operator of the Choquet integral. Then, the POP Choquet-like integrals are defined to perform the combinatorial step of ensembles of classifiers and to generalize the GN modularity for the fuzzy community network. Finally, two new algorithms are designed for experiments, and some computational experiments with other algorithms show the importance of POP Choquet-like integrals. All of the experimental results show that our algorithms are practical.

Keywords: data analysis; pseudo overlap function; Choquet-like integral; classifier ensemble; community network detection



Citation: Zhang, X.; Jiang, H.; Wang, J. New Classifier Ensemble and Fuzzy Community Detection Methods Using POP Choquet-like Integrals. *Fractal Fract.* **2023**, *7*, 588. <https://doi.org/10.3390/fractalfract7080588>

Academic Editor: Carlo Cattani

Received: 17 June 2023

Revised: 27 July 2023

Accepted: 28 July 2023

Published: 30 July 2023



Copyright: © 2023 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

A decade ago, data seemed primarily the preserve of researchers and a few professional managers. It has entered our everyday lexicon with the constant refinement and iteration of hardware and software, and our digital world generates and consumes vast amounts of data daily. Data analysis topics include classification, clustering, mining, data association analysis, etc. This paper mainly studies data classification and clustering algorithms in data analysis.

1.1. Classifier Ensembles and Choquet Integrals

Data classification is a model that describes and distinguishes data classes based on existing data sets. Usually, such a model is also called a classifier. The commonly used classifier algorithms include the Naive Bayes method [1], support vector machine [2], decision tree [3], artificial neural network [4], k-nearest neighbor [5], etc. At the same time, in real life, a single classification algorithm can only handle some specific types of data, which makes the classification work tedious. Therefore, people are eager to implement a classification algorithm that can deal with various types of data, and classifier ensembles can do this work well.

An ensemble of classifiers is a mechanism that fuses multiple classifiers, which consists of two steps: classification and fusion. It combines the benefits of multiple classifiers so that

any type of data set can be classified well. It also reduces the error generated by a single classifier, namely overfitting. Recently, in the literature, classifier ensemble has been used in areas such as disease detection [6], social networks [7] and mood recognition [8]. In these articles, the authors adopted a classifier ensemble for the current context, which yielded good results. However, they simply replaced or improved the base classifier algorithm of classifier ensembles and chose a simple approach to the fusion process, such as voting and weighting.

Obviously, using the Choquet integral as a fusion operator is more complex, and many scholars have used this ensemble approach to study in different contexts. In the literature, the author of [9] used the Choquet integral to fuse multiple classifiers to design credit score models. The results show that this model improves the accuracy of the classification algorithm. In [8], the author used the Choquet integral to integrate the segmentation results of multiple classifiers to obtain a better image segmentation scheme. In [10], the author used a classifier ensemble based on the Choquet integral to classify malware in smartphones and experimentally verified that the method is superior to a single classifier, with an accuracy of 95.08%. Especially, in [11], Batista et al. thought that it was more appropriate to replace the product operator with the same property in the Choquet integral with a quasi-overlap function that did not require associativity and continuity, so the Choquet-like integral based on quasi-overlap functions (CQO integral) was constructed. The author used the CQO integral to solve the classifier ensemble problem and compared it with other ensemble algorithms, and proved that it has advantages in solving the classifier ensemble problem.

Inspired by the CQO integral in [11] which is shown as

$$C_m^O(\vec{I}) = \sum_{i=1}^n \left(O(l_{(i)}, m(A_{(i)})) - O(l_{(i-1)}, m(A_{(i)})) \right),$$

we find the following two urgent tasks:

- (1) the CQO integral did not map to the [0,1] interval, similarly to many fusion functions, but rather to the [0,n] interval;
- (2) the CQO integral used the same O before and after replacing the product in the Choquet integral, which may lead to inconspicuous results for different inputs.

Therefore, this paper aims to design a new Choquet-like integral that not only naturally maps to the [0,1] interval but also replaces the product operator of the Choquet integral with two different pseudo overlap functions, which we named the pseudo overlap pair (POP) Choquet-like integral. We face two important questions for that: does the POP Choquet-like integral meet the boundary conditions and contacting (pre)aggregate functions? Does the POP Choquet-like integral still play a good role in classifier ensemble?

To solve these problems, we have certain requirements for selecting a pseudo-function pair, and some theories prove that the POP Choquet-like integral is a pre-aggregate function under some conditions. Moreover, we will use the POP Choquet-like integral as the fusion function for the classifiers ensemble and design an algorithm for experimental verification.

1.2. Community Network Detection and Choquet Integrals

Scholar Jain pointed out in [12] that clustering is classifying patterns (observations, data items, or feature vectors) without supervision. There are many clustering methods, and the common ones are nearest neighbor clustering [13]; fuzzy clustering [14]; artificial neural network clustering [15], etc. All of the above clustering methods can be used for complex network detection.

Complex network detection has been a popular topic in recent years because it involves a wide and deep range, including but not limited to telecommunications networks [16], community networks [17], traffic networks [18], and biological networks [19]. Among them, the nodes of the community networks are usually composed of people, and the connection between the nodes is dominated by familiarity, emotion, information, and other factors. Generally speaking, a community in a network is a subgraph characterized by close

connections between nodes within the subgraph but sparse connections between subgraphs. The study of community structure is significant in detecting community networks.

Considering the overlapping character of the complex community networks, the fuzzy community detection (FCD) methods were designed to calculate the membership degree of nodes in each community network. Many FCD methods previously proposed are only partially suitable for fuzzy frameworks, such as [20,21]. Specifically, in [22], the author used the overlap and the grouping functions to aggregate the membership of nodes to classes, generalized the classical GN modularity, and designed a community network detection algorithm that is completely suitable for fuzzy frameworks. However, when designing this algorithm, the author still selected the maximum as a group function; using the maximum function as an average could limit the results. Using fusion functions with non-average properties in most applications is more appropriate.

Generalized Choquet integrals are a class of non-average sum vector aggregation functions with good performance. Therefore, in this paper, the POP Choquet-like integral is used to replace the maximum operation in the definition of modularity in [22], which improves the membership processing of nodes after fuzzy clustering and can effectively detect overlapping communities. Moreover, we replace the overlap function in [22] with a pseudo overlap function that does not require commutativity, which is more in line with the actual situation. Then we will design a community network detection algorithm to prove our FCD method is effective.

1.3. Organizational Structure of This Paper

The rest of this article is framed as follows: Section 2 reviews pseudo overlap functions, Choquet integral and its generalizations, and preaggregate function. In Section 3, we introduce the pseudo overlap function pair and define the POP Choquet-like integral, discussing its properties in some cases. In Sections 4 and 5, we design a new ensemble classification algorithm and a new community network detection algorithm by using the defined POP Choquet-like integrals and design some experiments to prove their performance. Finally, Section 6 summarizes the full paper.

2. Preliminaries

This section aims to introduce the basic theories necessary for this article.

Definition 1 ([23]). Let $\vec{l} = (l_1, \dots, l_n)$ be a non-zero real vector of n -dimension. If for any $\vec{u} = (u_1, \dots, u_n) \in [0, 1]^n$ and $c > 0$, $\vec{u} + c\vec{l} = (u_1 + cl_1, \dots, u_n + cl_n) \in [0, 1]^n$, the function $Z : [0, 1]^n \rightarrow [0, 1]$ satisfies $Z(u_1 + cl_1, \dots, u_n + cl_n) > Z(u_1, \dots, u_n)$, then Z is \vec{l} -increasing.

Definition 2 ([23]). PAF : $[0, 1]^n \rightarrow [0, 1]$ is said to be an n -dimensional pre-aggregate function if it satisfies the all following conditions.

- (PAF1) directionally l -increasing: there exists a non-zero vector $\vec{l} = (l_1, \dots, l_n) \in [0, 1]^n$ such that the function PAF is \vec{l} increasing;
- (PAF2) $PAF(0, \dots, 0) = 0$;
- (PAF3) $PAF(1, \dots, 1) = 1$.

Definition 3 ([24]). A pseudo overlap function $PO : [0, 1]^2 \rightarrow [0, 1]$ is a binary function that satisfies all of the following properties:

- (PO1) $PO(x, y) = 0$ if and only if $xy = 0$;
- (PO2) $PO(x, y) = 1$ if and only if $xy = 1$;
- (PO3) PO is incremental;
- (PO4) PO is continuous.

In Table 1, we give some examples of pseudo overlap functions.

Table 1. Pseudo overlap functions.

Sequence Number	Nomenclature	Definition
1	PO_α	$PO_\alpha(x, y) = xy.$
2	PO_β	$PO_\beta(x, y) = \begin{cases} \frac{2xy}{x+y}, x + y \neq 0, \\ 0, x + y = 0. \end{cases}$
3	PO_χ	$PO_\chi(x, y) = \frac{2xy}{1+xy}.$
4	PO_δ	$PO_\delta(x, y) = \max\{\min\{x, \frac{y}{2}\}, x + y - 1\}.$
5	PO_ε	$PO_\varepsilon(x, y) = 0 \cdot 1xy^2 + 0.9 \max\{0, x + y - 1\}.$
6	PO_ϕ	$PO_\phi(x, y) = x^2y - 0.5x^2y(1 - x)(1 - y).$
7	PO_φ	$PO_\varphi(x, y) = \min\{x^2, y^2\}$
8	PO_γ	$PO_\gamma(x, y) = \begin{cases} \frac{2xy}{0.5x+1.5y}, 0.5x + 1.5y \neq 0, \\ 0, \text{others}. \end{cases}$
9	PO_η	$PO_\eta(x, y) = \frac{2x^2y^2}{1+x^2y^2}.$
10	PO_t	$PO_t(x, y) = \begin{cases} \frac{xy}{0.5x+0.5y}, 0.5x + 0.5y \neq 0, \\ 0, \text{others}. \end{cases}$
11	PO_κ	$PO_\kappa(x, y) = \begin{cases} \frac{3xy}{x+2y}, x + 2y \neq 0, \\ 0, \text{others}. \end{cases}$
12	PO_λ	$PO_\lambda(x, y) = \frac{2x^{10}y^{10}}{1+x^{10}y^{10}}$
13	PO_μ	$PO_\mu(x, y) = x^2y + x^2y(1 - x)(1 - y)$
14	PO_ν	$PO_\nu(x, y) = \min\{x^3, y^3\}$
15	PO_π	$PO_\pi(x, y) = xy^2$

Definition 4 ([25]). Let $N = \{1, 2, \dots, n\}$. $\forall X \subseteq Y \subseteq N$. If the following conditions are met, the function $m : P(N) \rightarrow [0, 1]$ will be called a fuzzy measure.

- (m1) $m(\emptyset) = 0;$
- (m2) $m(N) = 1;$
- (m3) Incremental: $m(X) \leq m(Y).$

Definition 5 ([25]). Let $N = \{1, 2, \dots, n\}$, $H \subseteq N$. The most classical fuzzy measure is the uniform fuzzy measure, which is defined as follows:

$$m_U(H) = \frac{|H|}{n}.$$

Some of the functions in Table 1 are of the same family, such as pseudo-overlapping functions with serial numbers 2 and 8, where the arguments α are 1 and 0.5 for pseudo-overlapping function family

$$PO(x, y) = \begin{cases} \frac{2xy}{\alpha x + (2-\alpha)y}, \alpha x + (2-\alpha)y \neq 0, & \alpha \in (0, 2), \\ 0, \alpha x + (2-\alpha)y = 0. \end{cases}$$

respectively. Due to the subsequent need to distinguish the size relationships between functions, it is necessary to determine the parameters of each family of functions. After many experiments, the pseudo-overlapping functions of these deterministic parameters in Table 1 are best fused in the two algorithms in this paper.

Another classical fuzzy measure is the most commonly used one, the g_λ fuzzy measure.

Definition 6 ([26]). Let $\lambda \geq -1$, $N = \{1, 2, \dots, n\}$. Function $m : P(N) \rightarrow [0, 1]$ satisfies the g_λ law that for any disjoint sequence $\{E_1, E_2, \dots, E_n, \dots\}$ in $P(N)$, and their union is also in $P(N)$, there is

$$m\left(\bigcup_{i=1}^{\infty} E_i\right) = \begin{cases} \frac{1}{\lambda} \left\{ \prod_{i=1}^{\infty} [1 + \lambda m(E_i)] - 1 \right\}, \lambda \neq 0, \\ \sum_{i=1}^{\infty} m(E_i), \lambda = 0. \end{cases} \tag{1}$$

Here, the λ coefficient is obtained by solving the following equation:

$$\lambda + 1 = \prod_{x \in A} (1 + \lambda m(\{x\})).$$

The formula (1) is called g_λ fuzzy measure. When g_λ fuzzy measure is used in our experiment, each $x \in E_i$ initial measure value $m(\{x\})$ is determined by the membership value in the initial classification result.

Definition 7 ([25]). Let $m : P(N) \rightarrow [0, 1]$ be a fuzzy measure. For any $\vec{l} = (l_1, l_2, \dots, l_n) \in [0, 1]^n$, the discrete Choquet integral $C_m : [0, 1]^n \rightarrow [0, 1]$ is constructed as follows:

$$C_m(\vec{l}) = \sum_{i=1}^n \left((l_{(i)} \cdot m(A_{(i)})) - (l_{(i-1)} \cdot m(A_{(i)})) \right),$$

where $(l_{(1)}, \dots, l_{(n)})$ is an increasing permutation of \vec{l} , i.e., $0 \leq l_{(1)} \leq \dots \leq l_{(n)}$, and the initial value $l_{(0)} = 0$. $A_{(i)} = \{(i), \dots, (n)\}$ is the subset of indices corresponding to the $n - i + 1$ largest components of \vec{l} .

Lucca et al. [23] constructed a series of Choquet-like integrals by replacing the product operators of Choquet integrals with t-norm, copula, and fusion function pair, which are proven to have good performance in fuzzy-rule classification systems (FRBCS). In this paper, Lucca's recently-constructed Choquet-like integral based on pair of fusion functions is referred to as $C_m^{(F_1, F_2)}$. The definition of $C_m^{(F_1, F_2)}$ is as follows:

Definition 8 ([23]). Let $m : P(N) \rightarrow [0, 1]$ be a fuzzy measure, (F_1, F_2) be a pair of fusion functions satisfying $\forall x, y \in [0, 1], F_1(x, y) \geq F_2(x, y)$. For any $\vec{l} = (l_1, l_2, \dots, l_n) \in [0, 1]^n$, the integral $C_m^{(F_1, F_2)} : [0, 1]^n \rightarrow [0, 1]$ is constructed as follows:

$$C_m^{(F_1, F_2)}(\vec{l}) = \min \left\{ 1, l_{(1)} + \sum_{i=2}^n \left(F_1(l_{(i)}, m(A_{(i)})) - F_2(l_{(i-1)}, m(A_{(i)})) \right) \right\},$$

where $(l_{(1)}, \dots, l_{(n)})$ is an increasing permutation of \vec{l} , i.e., $0 \leq l_{(1)} \leq \dots \leq l_{(n)}$, and the initial value $l_{(0)} = 0$. $A_{(i)} = \{(i), \dots, (n)\}$ is the subset of indices corresponding to the $n - i + 1$ largest components of \vec{l} .

To improve the quality of the classifier ensemble, in [11], the CQO integral is used as the fusion function in an ensemble. The definition of CQO is as follows

Definition 9 ([11]). Let $m : P(N) \rightarrow [0, 1]$ be a fuzzy measure, $O : [0, 1]^2 \rightarrow [0, 1]$ be a quasi overlap function. for any $\vec{l} = (l_1, l_2, \dots, l_n) \in [0, 1]^n$, the CQO integral $C_m^O : [0, 1]^n \rightarrow [0, n]$ is constructed as follows:

$$C_m^O(\vec{l}) = \sum_{i=1}^n \left(O(l_{(i)}, m(A_{(i)})) - O(l_{(i-1)}, m(A_{(i)})) \right),$$

where $(l_{(1)}, \dots, l_{(n)})$ is an increasing permutation of \vec{l} , i.e., $0 \leq l_{(1)} \leq \dots \leq l_{(n)}$, and the initial value $l_{(0)} = 0$. $A_{(i)} = \{(i), \dots, (n)\}$ is the set of index that corresponds to the previous $n - i + 1$ largest element.

3. POP Choquet-like Integral

Although several Choquet-like integrals have been proposed for various scenarios, their structures have always been somewhat unnatural. For example, $C_m^{(F_1, F_2)}$ directly maps integral values greater than 1 to 1, making it impossible to compare the results. In addition, the integral C_m^O maps the value to $[0, n]$. However, the codomain of the common fusion

function is $[0, 1]$, so it needs some restrictions to reduce its range to $[0, 1]$, which makes the application more difficult.

In addition, since the product operator in the original Choquet integral does not force commutativity and associativity, the pseudo overlap function is a good substitution function. Additionally, the non-average function is more competitive than the average function in many application scenarios. Therefore, in this section, we use pseudo overlap function pair (PO_1, PO_2) to generalize the Choquet integral to obtain the POP Choquet-like integral. The following is its construction procedure.

Definition 10. For given two pseudo overlap functions $PO_1, PO_2 : [0, 1]^2 \rightarrow [0, 1]$. (PO_1, PO_2) is called a pseudo overlap function pair as long as it satisfies $PO_1(x, y) \geq PO_2(x, y)$ for any $x, y \in [0, 1]$.

We compare the size of the 16 functions in Table 1 and list 15 pseudo-overlap function pairs in Table 2.

Table 2. The size relation of pseudo overlap functions in Table 1.

Sequence Number	Pseudo Overlap Function Pair	Expression
1	(PO_β, PO_α)	$PO_\beta(x, y) \geq PO_\alpha(x, y), \forall x, y \in [0, 1]$
2	(PO_β, PO_χ)	$PO_\beta(x, y) \geq PO_\chi(x, y), \forall x, y \in [0, 1]$
3	(PO_β, PO_δ)	$PO_\beta(x, y) \geq PO_\delta(x, y), \forall x, y \in [0, 1]$
4	(PO_β, PO_ϵ)	$PO_\beta(x, y) \geq PO_\epsilon(x, y), \forall x, y \in [0, 1]$
5	(PO_β, PO_ϕ)	$PO_\beta(x, y) \geq PO_\phi(x, y), \forall x, y \in [0, 1]$
6	(PO_χ, PO_λ)	$PO_\chi(x, y) \geq PO_\lambda(x, y), \forall x, y \in [0, 1]$
7	(PO_ϕ, PO_ν)	$PO_\phi(x, y) \geq PO_\nu(x, y), \forall x, y \in [0, 1]$
8	(PO_α, PO_ϕ)	$PO_\alpha(x, y) \geq PO_\phi(x, y), \forall x, y \in [0, 1]$
9	(PO_α, PO_ν)	$PO_\alpha(x, y) \geq PO_\nu(x, y), \forall x, y \in [0, 1]$
10	(PO_i, PO_α)	$PO_i(x, y) \geq PO_\alpha(x, y), \forall x, y \in [0, 1]$
11	(PO_κ, PO_α)	$PO_\kappa(x, y) \geq PO_\alpha(x, y), \forall x, y \in [0, 1]$
12	(PO_γ, PO_α)	$PO_\gamma(x, y) \geq PO_\alpha(x, y), \forall x, y \in [0, 1]$
13	(PO_γ, PO_δ)	$PO_\gamma(x, y) \geq PO_\delta(x, y), \forall x, y \in [0, 1]$
14	(PO_γ, PO_η)	$PO_\gamma(x, y) \geq PO_\eta(x, y), \forall x, y \in [0, 1]$
15	(PO_α, PO_π)	$PO_\alpha(x, y) \geq PO_\pi(x, y), \forall x, y \in [0, 1]$

For the n-dimensional incremental vector $\vec{l} = (l_1, l_2, \dots, l_n) \in [0, 1]^n$, since its components may contain duplicates, the order of components of the incremental vector \vec{l} can be controversial in constructing the POP Choquet-like integral afterward. Hence, vector \vec{l} needs to be reduced in dimension to ensure that it has no duplicate components.

Definition 11. For a given $\vec{l} = (l_1, l_2, \dots, l_n) \in [0, 1]^n (n \in N, n \geq 2)$. The dimensionality reduction function is defined as $DR : [0, 1]^n \rightarrow [0, 1] \cup [0, 1]^2 \cup \dots \cup [0, 1]^n$, in the form:

$$DR(l_1, l_2, \dots, l_n) = \vec{y} = (y_1, y_2, \dots, y_k), (1 \leq k \leq n)$$

which satisfies the following conditions:

- (DR1) $y_i < y_{i+1}, i = 1, 2, \dots, k - 1;$
- (DR2) $\forall y_i \in \vec{y}, \exists l_j, l_{j+1}, \dots, l_{j+p}, 1 \leq j \leq n, 0 \leq p \leq n - j, s.t. y_i = l_j = l_{j+1} = \dots = l_{j+p}.$

After defining the pseudo overlap function pair and dimensionality reduction function, we can design a POP Choquet-like integral. After that, we will also explore the boundary conditions for POP Choquet-like integral and the conditions under which the POP Choquet-like integral is a pre-aggregate function.

Definition 12. Let $N = \{1, 2, \dots, n\}$ be a finite set, $m : P(N) \rightarrow [0, 1]$ be a fuzzy measure, (PO_1, PO_2) be a pseudo overlap function pair that satisfying $\forall x, y \in [0, 1], PO_1(x, y) \geq PO_2(x, y)$, $DR : [0, 1]^n \rightarrow [0, 1] \cup [0, 1]^2 \cup \dots \cup [0, 1]^n$ be a dimensionality reduction function. For any $\vec{x} = (x_1, x_2, \dots, x_n) \in [0, 1]^n$, POP Choquet-like integral $C_m^{(PO_1, PO_2)} : [0, 1]^n \rightarrow [0, 1]$ is defined as follows:

$$C_m^{(PO_1, PO_2)}(\vec{x}) = \frac{\sum_{j=1}^k \left(PO_1(y_{(j)}, m(A_{(j)})) - PO_2(y_{(j-1)}, m(A_{(j)})) \right)}{k},$$

of which

$$\vec{y} = (y_1, y_2, \dots, y_k) = DR(x_1, x_2, \dots, x_n),$$

and $(y_{(1)}, \dots, y_{(k)})$ is an increasing permutation on \vec{y} , that is, $0 \leq y_{(1)} \leq \dots \leq y_{(k)}$, where $y_{(0)} = 0$ and $A_{(j)} = \{(j), \dots, (k)\}$ is the subset of indices corresponding to the $k - j + 1$ largest components of \vec{y} .

Table 3 lists fifteen POP Choquet-like integrals based on pseudo overlap function pairs.

Remark 1. $C_m^{(PO_1, PO_2)}$ is well defined for arbitrary pairs of pseudo overlap functions and fuzzy measures m .

Obtained by Definition 12, for any given $\vec{x} = (x_1, x_2, \dots, x_n) \in [0, 1]^n$, dimensionality reduction function

$$DR(x_1, x_2, \dots, x_n) = (y_1, y_2, \dots, y_k).$$

and $(y_{(1)}, \dots, y_{(k)})$ is an increasing permutation on \vec{y} , so

$$\begin{aligned} C_m^{(PO_1, PO_2)}(x_1, x_2, \dots, x_n) &= \frac{\sum_{j=1}^k \left(PO_1(y_{(j)}, m(A_{(j)})) - PO_2(y_{(j-1)}, m(A_{(j)})) \right)}{k} \\ &\geq \frac{\sum_{j=1}^k \left(PO_1(y_{(j)}, m(A_{(j)})) - PO_1(y_{(j-1)}, m(A_{(j)})) \right)}{k} \\ &\geq 0. \\ C_m^{(PO_1, PO_2)}(x_1, x_2, \dots, x_n) &= \frac{\sum_{j=1}^k \left(PO_1(y_{(j)}, m(A_{(j)})) - PO_2(y_{(j-1)}, m(A_{(j)})) \right)}{k} \\ &\leq \frac{\sum_{j=1}^k PO_1(y_{(j)}, m(A_{(j)}))}{k} \\ &\leq 1. \end{aligned}$$

Proposition 1. Let (PO_1, PO_2) be a pair of pseudo overlap functions, m be a fuzzy measure, DR be a dimensionality reduction function. We have $C_m^{(PO_1, PO_2)}(0, 0, \dots, 0) = 0$, and $C_m^{(PO_1, PO_2)}(1, 1, \dots, 1) = 1$.

Proof.

$$\begin{aligned} C_m^{(PO_1, PO_2)}(0, 0, \dots, 0) &= C_m^{(PO_1, PO_2)}(0) \\ &= PO_1(0, m(A_{(1)})) - 0 \\ &= 0. \end{aligned}$$

$$\begin{aligned} C_m^{(PO_1, PO_2)}(1, 1, \dots, 1) &= C_m^{(PO_1, PO_2)}(1) \\ &= PO_1(1, m(A_{(1)})) - 0 \\ &= PO_1(1, 1) \\ &= 1. \end{aligned}$$

□

Table 3. Example of POP Choquet-like integrals.

Sequence Number	POP Integral	Expression
1	$C_m^{(PO_\beta, PO_\alpha)}$	$C_m^{(PO_\beta, PO_\alpha)}(\vec{x}) = \begin{cases} \frac{\sum_{j=1}^k \left(\frac{2y_{(j)} \cdot m(A_{(j)})}{y_{(j)} + m(A_{(j)})} - y_{(j-1)} \cdot m(A_{(j)}) \right)}{k}, & y_{(j)} + m(A_{(j)}) \neq 0 \\ 0, & y_{(j)} + m(A_{(j)}) = 0 \end{cases}$
2	$C_m^{(PO_\beta, PO_\chi)}$	$C_m^{(PO_\beta, PO_\chi)}(\vec{x}) = \begin{cases} \frac{\sum_{j=1}^k \left(\frac{2y_{(j)} \cdot m(A_{(j)})}{y_{(j)} + m(A_{(j)})} - \frac{2y_{(j-1)} \cdot m(A_{(j)})}{1 + y_{(j-1)} \cdot m(A_{(j)})} \right)}{k}, & y_{(j)} + m(A_{(j)}) \neq 0 \\ 0, & y_{(j)} + m(A_{(j)}) = 0 \end{cases}$
3	$C_m^{(PO_\beta, PO_\delta)}$	$C_m^{(PO_\beta, PO_\delta)}(\vec{x}) = \begin{cases} \frac{\sum_{j=1}^k \left(\frac{2y_{(j)} \cdot m(A_{(j)})}{y_{(j)} + m(A_{(j)})} - \max\left\{ \min\left\{ y_{(j-1)}, \frac{m(A_{(j)})}{2} \right\} \right\} \cdot y_{(j-1)} + m(A_{(j)}) - 1 \right)}{k}, & y_{(j)} + m(A_{(j)}) \neq 0 \\ 0, & y_{(j)} + m(A_{(j)}) = 0 \end{cases}$
4	$C_m^{(PO_\beta, PO_\epsilon)}$	$C_m^{(PO_\beta, PO_\epsilon)}(\vec{x}) = \begin{cases} \frac{\sum_{j=1}^k \left(\frac{2y_{(j)} \cdot m(A_{(j)})}{y_{(j)} + m(A_{(j)})} - 0.1 \cdot y_{(j-1)} \cdot m(A_{(j)})^2 + 0.9 \cdot \max\{0, y_{(j-1)} + m(A_{(j)}) - 1\} \right)}{k}, & y_{(j)} + m(A_{(j)}) \neq 0 \\ 0, & y_{(j)} + m(A_{(j)}) = 0 \end{cases}$
5	$C_m^{(PO_\beta, PO_\phi)}$	$C_m^{(PO_\beta, PO_\phi)}(\vec{x}) = \begin{cases} \frac{\sum_{j=1}^k \left(\frac{2y_{(j)} \cdot m(A_{(j)})}{y_{(j)} + m(A_{(j)})} - y_{(j-1)}^2 \cdot m(A_{(j)}) + 0.5 \cdot y_{(j-1)}^2 \cdot m(A_{(j)}) \cdot (1 - y_{(j-1)}) \cdot (1 - m(A_{(j)})) \right)}{k}, & y_{(j)} + m(A_{(j)}) \neq 0 \\ 0, & y_{(j)} + m(A_{(j)}) = 0 \end{cases}$
6	$C_m^{(PO_\chi, PO_\lambda)}$	$C_m^{(PO_\chi, PO_\lambda)}(\vec{x}) = \frac{\sum_{j=1}^k \left(\frac{2y_{(j)} \cdot m(A_{(j)})}{1 + y_{(j)} \cdot m(A_{(j)})} - \frac{2y_{(j-1)}^{10} \cdot m(A_{(j)})^{10}}{1 + y_{(j-1)}^{10} \cdot m(A_{(j)})^{10}} \right)}{k}$
7	$C_m^{(PO_\alpha, PO_\nu)}$	$C_m^{(PO_\alpha, PO_\nu)}(\vec{x}) = \frac{\sum_{j=1}^k \left(y_{(j)} \cdot m(A_{(j)}) - \min\{y_{(j-1)}^3, m^3(A_{(j)})\} \right)}{k}$
8	$C_m^{(PO_\phi, PO_\nu)}$	$C_m^{(PO_\phi, PO_\nu)}(\vec{x}) = \frac{\sum_{j=1}^k \left(\min\{y_{(j)}^2, m^2(A_{(j)})\} - \min\{y_{(j-1)}^2, m^2(A_{(j)})\} \right)}{k}$
9	$C_m^{(PO_\alpha, PO_\phi)}$	$C_m^{(PO_\alpha, PO_\phi)}(\vec{x}) = \frac{\sum_{j=1}^k \left(y_{(j)} \cdot m(A_{(j)}) - \min\{y_{(j-1)}^2, m^2(A_{(j)})\} \right)}{k}$
10	$C_m^{(PO_\gamma, PO_\alpha)}$	$C_m^{(PO_\gamma, PO_\alpha)}(\vec{x}) = \begin{cases} \frac{\sum_{j=1}^k \left(\frac{y_{(j)} \cdot m(A_{(j)})}{0.5 \cdot (y_{(j)} + m(A_{(j)}))} - y_{(j-1)} \cdot m(A_{(j)}) \right)}{k}, & y_{(j)} + m(A_{(j)}) \neq 0 \\ 0, & y_{(j)} + m(A_{(j)}) = 0 \end{cases}$
11	$C_m^{(PO_\alpha, PO_\alpha)}$	$C_m^{(PO_\alpha, PO_\alpha)}(\vec{x}) = \begin{cases} \frac{\sum_{j=1}^k \left(\frac{3y_{(j)} \cdot m(A_{(j)})}{y_{(j)} + 2 \cdot m(A_{(j)})} - y_{(j-1)} \cdot m(A_{(j)}) \right)}{k}, & y_{(j)} + 2 \cdot m(A_{(j)}) \neq 0 \\ 0, & y_{(j)} + 2 \cdot m(A_{(j)}) = 0 \end{cases}$
12	$C_m^{(PO_\gamma, PO_\alpha)}$	$C_m^{(PO_\gamma, PO_\alpha)}(\vec{x}) = \begin{cases} \frac{\sum_{j=1}^k \left(\frac{2y_{(j)} \cdot m(A_{(j)})}{0.5 \cdot y_{(j)} + 1.5 \cdot m(A_{(j)})} - y_{(j-1)} \cdot m(A_{(j)}) \right)}{k}, & 0.5 \cdot y_{(j)} + 1.5 \cdot m(A_{(j)}) \neq 0 \\ 0, & 0.5 \cdot y_{(j)} + 1.5 \cdot m(A_{(j)}) = 0 \end{cases}$
13	$C_m^{(PO_\gamma, PO_\delta)}$	$C_m^{(PO_\gamma, PO_\delta)}(\vec{x}) = \begin{cases} \frac{\sum_{j=1}^k \left(\frac{2y_{(j)} \cdot m(A_{(j)})}{0.5 \cdot y_{(j)} + 1.5 \cdot m(A_{(j)})} - \max\left\{ \min\left\{ y_{(j-1)}, \frac{m(A_{(j)})}{2} \right\} \right\} \cdot y_{(j-1)} + m(A_{(j)}) - 1 \right)}{k}, & 0.5 \cdot y_{(j)} + 1.5 \cdot m(A_{(j)}) \neq 0 \\ 0, & 0.5 \cdot y_{(j)} + 1.5 \cdot m(A_{(j)}) = 0 \end{cases}$
14	$C_m^{(PO_\gamma, PO_\eta)}$	$C_m^{(PO_\gamma, PO_\eta)}(\vec{x}) = \begin{cases} \frac{\sum_{j=1}^k \left(\frac{2y_{(j)} \cdot m(A_{(j)})}{0.5 \cdot y_{(j)} + 1.5 \cdot m(A_{(j)})} - \frac{2y_{(j-1)}^2 \cdot m^2(A_{(j)})}{1 + y_{(j-1)}^2 \cdot m^2(A_{(j)})} \right)}{k}, & 0.5 \cdot y_{(j)} + 1.5 \cdot m(A_{(j)}) \neq 0 \\ 0, & 0.5 \cdot y_{(j)} + 1.5 \cdot m(A_{(j)}) = 0 \end{cases}$
15	$C_m^{(PO_\alpha, PO_\pi)}$	$C_m^{(PO_\alpha, PO_\pi)}(\vec{x}) = \frac{\sum_{j=1}^k \left(y_{(j)} \cdot m(A_{(j)}) - y_{(j-1)}^5 \cdot m(A_{(j)})^5 \right)}{k}$

Proposition 2. Let (PO_1, PO_2) be a pair of pseudo overlap functions, m be a fuzzy measure, and DR be a dimensionality reduction function. $C_m^{(PO_1, PO_2)}$ is idempotent if and only if PO_1 has a neutral element.

Proof. Let $l \in [0, 1]$, then

$$\begin{aligned} C_m^{(PO_1, PO_2)}(l, l, \dots, l) &= C_m^{(PO_1, PO_2)}(l) \\ &= PO_1(l, m(A_{(1)})) - PO_2(0, m(A_{(1)})) \\ &= PO_1(l, m(A_{(1)})) \\ &= PO_1(l, 1). \end{aligned}$$

So $C_m^{(PO_1, PO_2)}$ is idempotent if and only if $PO_1(l, 1) = l$. \square

Example 1. POP Choquet-like integral $C_m^{(PO_\alpha, PO_\phi)}$ is idempotent because $PO_\alpha(x, y) = xy$ satisfies $PO_\alpha(x, 1) = x$.

Proposition 3. Let (PO_1, PO_2) be a pair of pseudo overlap functions and m be a fuzzy measure. $C_m^{(PO_1, PO_2)}$ is a pre-aggregate function if the following conditions are true:

- (1) $\frac{\partial PO_1(v,w)}{\partial v} \geq \frac{\partial PO_2(u,w)}{\partial u}, \forall u, v, w \in [0, 1], u < v;$
- (2) $PO_1(ku, v) = kPO_1(u, v), PO_2(ku, v) = kPO_2(u, v), \forall x, y \in [0, 1]$ and $k \in (0, 1]$.

Proof. We only need to prove $C_m^{(PO_1, PO_2)}$ is $\vec{1} = (1, \dots, 1)$ -increasing, that is, for each $u, v, w, c, u + c, v + c \in [0, 1]$ and $u < v, PO_1(v, w) - PO_2(u, w) \leq PO_1(v + c, w) - PO_2(u + c, w)$ is established.

From condition 1, for each sufficiently small c , we can obtain the following:

$$\frac{PO_1(v + c, w) - PO_1(v, w)}{c} \geq \frac{PO_2(u + c, w) - PO_2(u, w)}{c},$$

$$PO_1(v + c, w) - PO_1(v, w) \geq PO_2(u + c, w) - PO_2(u, w),$$

$$PO_1(v + c, w) - PO_2(u + c, w) \geq PO_1(v, w) - PO_2(u, w).$$

From condition 2, for each not sufficiently small c , always exist $k \in (0, 1]$ makes $c_1 = c/k$ sufficiently small and greater than 0, then we have

$$\frac{kPO_1(v/k + c_1, w) - kPO_1(v/k, w)}{c_1} \geq \frac{kPO_2(u/k + c_1, w) - kPO_2(u/k, w)}{c_1},$$

$$PO_1(v/k + c_1, w) - PO_1(v/k, w) \geq PO_2(u/k + c_1, w) - PO_2(u/k, w),$$

think of $v/k + c_1$ as $v + c, u/k + c_1$ as $u + c$, then we have

$$PO_1(v + c, w) - PO_1(v, w) \geq PO_2(u + c, w) - PO_2(u, w),$$

$$PO_1(v + c, w) - PO_2(u + c, w) \geq PO_1(v, w) - PO_2(u, w).$$

Therefore, for each $u, v, w, c, u + c, v + c \in [0, 1]$, and $u < v, PO_1(v, w) - PO_2(u, w) \leq PO_1(v + c, w) - PO_2(u + c, w)$ is established, $C_m^{(PO_1, PO_2)}$ is a pre-aggregate function. \square

Remark 2. Condition 2 of Proposition 3 is a special kind of homogeneity, similar to the homogeneity proposed in [27]. Specifically, according to the definition given in [27], when PO_1 and PO_2 are overlap functions, we have

$$PO_1(kx, ky) = kPO_1(x, ky)$$

$$= kPO_1(ky, x)$$

$$= k^2PO_1(y, x)$$

$$= k^2PO_1(x, y).$$

The same is true for PO_2 . Therefore, condition 2 of Proposition 3 is a special homogeneity.

Example 2. $C_m^{(PO_\alpha, PO_\pi)}$ is a pre-aggregate function because

- (1) $\frac{\partial PO_\alpha(x,y)}{\partial x} = y, \frac{\partial PO_\pi(x,y)}{\partial x} = y^2$ and $x, y \in [0, 1]$, so $\frac{\partial PO_\alpha(x,y)}{\partial x} \geq \frac{\partial PO_\pi(x,y)}{\partial x}$.
- (2) $PO_\alpha(kx, y) = kxy, PO_\pi(kx, y) = kxy^2$, satisfying the (2) of Proposition 3.

Similarly, all POP integrals $C_m^{(PO_1, PO_2)}$ that satisfy $PO_1 = PO_2$ are pre-aggregate functions.

To illustrate the advantages of $C_m^{(PO_1, PO_2)}$ proposed by us, an example is given below to preliminarily compare it with integral $C_m^{(F_1, F_2)}$ and integral C_m^O . To facilitate calculation, the uniform fuzzy measure is selected.

Example 3. Given two randomly increasing arrays $\vec{a} = (0.6, 0.7, 0.8, 0.9)$ and $\vec{b} = (0.8, 0.8, 0.9, 0.9)$. We take them as input and calculate them in $C_m^{(F_1, F_2)}$, C_m^O , and $C_m^{(PO_1, PO_2)}$, and analyze the results.

First, to make the comparison easier, select integral $C_m^{(GM, TL)}$ in [23], C_m^{OC} in [11], and $C_m^{(PO_\chi, PO_\lambda)}$ defined in Section 3. It is easy to obtain the fuzzy measure value of \vec{a} as (1.0, 0.75, 0.5, 0.25). For input, the three integrals calculation steps are as follows:

$$C_m^{(GM, TL)}(\vec{a}) = \min \left\{ \begin{array}{l} 1, GM(0.6, 1.0) + GM(0.7, 0.75) - TL(0.6, 0.75) + GM(0.8, 0.5) \\ -TL(0.7, 0.5) + GM(0.9, 0.25) - TL(0.8, 0.25) \end{array} \right\}$$

$$= \min\{1, 2.01\}$$

$$= 1.$$

$$C_m^{OC}(\vec{a}) = O_C(0.6, 1.0) + O_C(0.7, 0.75) - O_C(0.6, 0.75) + O_C(0.8, 0.5)$$

$$- O_C(0.7, 0.5) + O_C(0.9, 0.25) - O_C(0.8, 0.25)$$

$$= 0.70.$$

$$C_m^{(PO_\chi, PO_\lambda)}(\vec{a}) = \frac{PO_\chi(0.6, 1.0) + PO_\chi(0.7, 0.75) - PO_\lambda(0.6, 0.75) + PO_\chi(0.8, 0.5)}{4}$$

$$- PO_\lambda(0.7, 0.5) + PO_\chi(0.9, 0.25) - PO_\lambda(0.8, 0.25)$$

$$\approx 0.59.$$

For input \vec{b} , in the integral C_m^{OC} , its fuzzy measure value still is (1.0, 0.75, 0.5, 0.25). In the integral $C_m^{(GM, TL)}$ and the integral $C_m^{(PO_\chi, PO_\lambda)}$, its fuzzy measure value becomes [1.0, 0.5] because of the function of dimensionality reduction function. For input \vec{b} , the three integrals calculation steps are as follows:

$$C_m^{(GM, TL)}(\vec{b}) = \min\{1, GM(0.8, 1.0) + GM(0.9, 0.5) - TL(0.8, 0.5)\}$$

$$= \min\{1, 1.265\}$$

$$= 1.$$

$$C_m^{OC}(\vec{b}) = O_C(0.8, 1.0) + O_C(0.8, 0.75) - O_C(0.8, 0.75) + O_C(0.9, 0.5)$$

$$- O_C(0.8, 0.5) + O_C(0.9, 0.25) - O_C(0.9, 0.25)$$

$$= O_C(0.8, 1.0) + O_C(0.9, 0.5) - O_C(0.8, 0.5)$$

$$= 0.78.$$

$$C_m^{(PO_\chi, PO_\lambda)}(\vec{b}) = \frac{PO_\chi(0.8, 1.0) + PO_\chi(0.9, 0.5) - PO_\lambda(0.8, 0.5)}{2}$$

$$\approx 0.75.$$

As you can see, whether we input \vec{a} or \vec{b} , the integral $C_m^{(GM, TL)}$ will compute 1, which is inappropriate. Additionally, it is easier to see the change in integral $C_m^{(PO_\chi, PO_\lambda)}$ large than in integral C_m^{OC} when the input changes.

4. Ensemble Algorithm Based on POP Choquet-like Integrals

A classifier ensemble is a means of fusing multiple classifiers' classification results. It can compensate with other classifiers when one classifier classification is not effective.

Classifier ensemble can be viewed as a two-tier pattern recognition structure, as shown in Figure 1. The first layer selects multiple classifiers as base classifiers, in which all base classifiers accept input conditions and output separately. This paper will give each base classifier the same training set as input. The second layer is the fusion method, which receives the output of multiple base classifiers, fuses them with a given fusion operator, and finally obtains a clear decision.

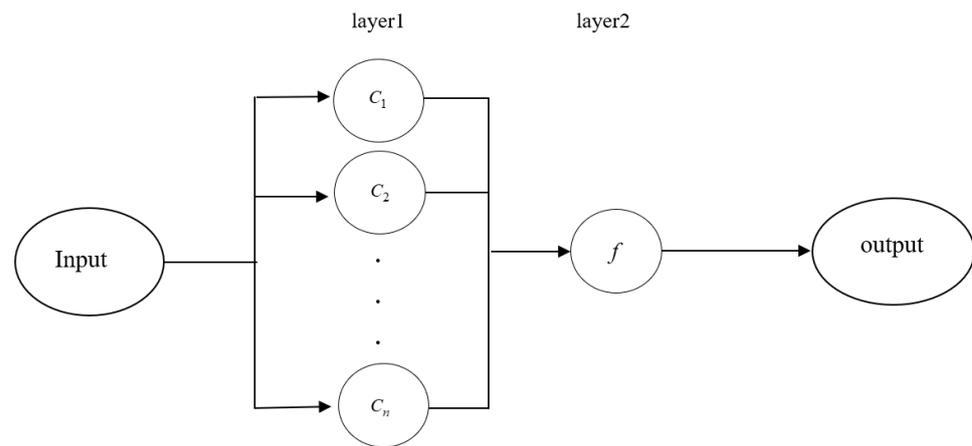


Figure 1. General architecture of classifier ensemble.

This paper uses the POP Choquet-like integral proposed in Section 3 as the fusion operator. This strategy is described in more detail below.

4.1. Algorithmic Framework

In the experiment of this algorithm, we test the pairs of pseudo-overlap functions by transforming them to obtain the $C_m^{(PO_1, PO_2)}$ integral (Using the g_λ fuzzy measure) with the best fusion function. This process is described in detail below, and the overall algorithmic framework is given.

Let H be a set of a given group of classifiers I_1, I_2, \dots, I_n , $H = \{I_1, I_2, \dots, I_n\}$. The first step in the ensemble is to enter the same training set into the classifier I_1, I_2, \dots, I_n to train the respective models, then test these models with the same test set and obtain a classification membership matrix for each base classifier. Suppose the input $\vec{x} = (x_1, x_2, \dots, x_r)$ is classified into k classes, let $P = (P_1, \dots, P_j, \dots, P_n)$, where P_j is the classification membership matrix I_j corresponding to the classifier. The row of P_j represents each element of the test set, and the column's index represents the class in which it is located. Each element of P_j represents the extent to which the test item corresponding to the row belongs to the class corresponding to the current column.

Next is the most critical step in an ensemble. Take the fused operator $C_m^{(PO_1, PO_2)}$ proposed in this article to de-fuse the elements of the coherence of the ranks of each component matrix of vector $P = (P_1, \dots, P_j, \dots, P_n)$. Then we obtain a matrix M of size $r \times k$, where element $M_{r,k}$ represents the degree to which the r -th data item belongs to class k . Let F be a matrix of size $r \times 1$, and its element consists of an index of columns corresponding to the maximum value of each row of M , representing the final classification result of the r -th data item. The model framework of Algorithm 1 is constructed below to describe this process more clearly.

Algorithm 1: Pseudo-code for our proposed ensemble model

Input: A pseudo overlap function pair (PO_1, PO_2) , classifier set $H = \{I_1, I_2, \dots, I_n\}$.

Output: F .

- 1 $P = (P_1, P_2, \dots, P_n) \leftarrow$ A vector consisting of elements from a classification membership matrix generated by I_1, I_2, \dots, I_n ;
 - 2 **for** $i = 0 \rightarrow r$ **do**
 - 3 **for** $j = 0 \rightarrow k$ **do**
 - 4 compute $M[i][j] = C_m^{(PO_1, PO_2)}(P_1[i][j], P_2[i][j], \dots, P_n[i][j])$;
 - 5 $F[i] = \text{MaxIndex}(M[i])$;
 - 6 **return** F .
-

4.2. Experimental Framework

- Step 1. Selection of data sets

This paper will select the data sets in the UCI database (<https://archive.ics.uci.edu/ml/datasets.php> (accessed on 11 December 2022)). This algorithm is aimed at multiple categories of data, so 12 data sets with a classification greater than or equal to 3 will be selected, with the number of feature items ranging from a few to a few hundred, as shown in Table 4.

Table 4. Description of the datasets used in our ensemble.

Name of Dataset	Instances Number	Attributes Number	Classes Number
Iris (IR)	150	4	3
Balance Scale (BS)	625	4	3
Winequality-red (WR)	1599	11	6
Waveform (WF)	5000	21	3
Optical recognition (OR)	5620	64	10
Cnae-9 (C9)	1080	857	9
Wireless indoor locatization (WI)	2000	7	4
Splice junction Gene sequences (SG)	2552	61	3
Car evaluation (CE)	1728	6	4
Maternal health risk (MH)	1014	6	6
Winequality-white (WW)	4898	11	7
Page-blocks (PB)	5472	10	5

- Step 2. Experimental preprocessing

(1) Deleting Missing Items

Because some data sets with incomplete data elements make the algorithm fail, delete the entire row of those data elements.

(2) Ten-fold cross validation

Before the experiment, the data items for each data set category are divided into ten pieces on average, resulting in ten subsets so that each subset has a different category of data items. Each time a subset is used as a test set, and the remaining nine are trained as training sets until each subset is selected once as a test set. There are 15 pseudo-overlap function pairs selected in this article. Each data set runs at least 150 times to obtain the final result.

- Step 3. Base algorithms and experiment details

This paper selects four basic classifiers: Naive Bayes, KNN (Euclidean distance, $K = 10$), neural network (three layers), and CART. The four base classifiers act as the four elements of the classifier set in a 1:1:1:1 scale, the step that is conducted through Python's scikit-learn library. To better demonstrate the experiment's results, multiply the F1-Measure value by 100.

Definition 13. Set the accuracy rate as P , and recall rate as R . The F -Measure is the weighted harmonic average of P and R , and is defined as:

$$F_{\alpha} = \frac{(1 + \alpha^2) \cdot P \cdot R}{\alpha^2 \cdot (P + R)}.$$

When the parameter $\alpha = 1$, it is the most common F1-Measure, defined as:

$$F1 = \frac{2 \cdot P \cdot R}{P + R}.$$

The values of P and R are in the interval $[0,1]$.

4.3. Experimental Results and Analysis

This subsection demonstrates the experimental results of the classifier ensemble algorithm of POP Choquet-like integral, and then we compare and analyze this algorithm with other classification algorithms.

In the first experiment of this algorithm, we test the performance of POP Choquet-like integrals $C_m^{(PO_1, PO_2)}$ based on different pseudo-overlap function pairs in the classifier ensemble. For the POP Choquet-like integrals in Table 5, we use pseudo overlap function pairs constructed by two identical pseudo overlap functions. In contrast, for the POP Choquet-like integrals in Table 6, we use pseudo overlap function pairs constructed by two different pseudo overlap functions. Tables 5 and 6 measure the classifier ensemble’s performance in this paper. For each data set, the best-performing calculation schemes are shown in bold.

Table 5. Ensemble algorithm performance of POP Choquet-like integrals that choose pairs with two identical pseudo overlap functions.

Dataset Integral	IR	BS	WR	WF	OR	C9	WI	SG	CE	MH	WW	PB	Mean
$C_m^{(PO_\alpha, PO_\alpha)}$	95.97	85.73	45.52	84.13	95.51	93.92	97.53	99.19	84.35	67.40	42.46	94.62	82.19
$C_m^{(PO_\beta, PO_\beta)}$	95.97	85.73	48.36	84.30	95.19	93.64	97.43	99.00	83.75	70.36	43.96	94.52	82.68
$C_m^{(PO_\chi, PO_\chi)}$	95.97	85.73	50.42	84.10	95.15	93.87	97.12	99.38	84.07	72.71	44.11	94.45	83.09
$C_m^{(PO_\delta, PO_\delta)}$	96.64	85.73	43.52	84.06	95.31	93.34	97.43	98.97	85.03	67.86	43.08	94.21	82.10
$C_m^{(PO_\epsilon, PO_\epsilon)}$	97.31	86.65	44.53	83.23	96.21	94.03	97.51	99.72	88.47	72.00	42.88	94.62	83.10
$C_m^{(PO_\phi, PO_\phi)}$	95.97	86.65	43.16	84.28	96.53	94.04	97.74	99.41	86.25	70.22	42.72	94.74	82.64
$C_m^{(PO_\psi, PO_\psi)}$	95.97	86.65	41.21	82.70	95.22	94.14	97.52	99.03	82.88	68.90	41.14	94.40	80.81
$C_m^{(PO_\gamma, PO_\gamma)}$	95.97	85.62	47.55	84.57	95.59	94.01	97.58	99.62	85.69	73.63	43.69	94.43	83.16
$C_m^{(PO_\eta, PO_\eta)}$	95.97	86.65	40.66	83.98	96.24	93.67	97.74	99.31	85.72	65.61	42.05	94.69	81.86
$C_m^{(PO_\iota, PO_\iota)}$	95.97	85.62	48.36	84.30	95.19	93.54	97.43	99.00	83.75	70.36	43.96	94.52	82.67
$C_m^{(PO_\kappa, PO_\kappa)}$	95.97	85.62	48.21	84.33	95.41	94.10	97.43	99.31	85.46	71.25	43.53	94.40	82.92
$C_m^{(PO_\lambda, PO_\lambda)}$	95.97	85.62	42.31	83.92	95.34	93.58	97.12	99.05	82.98	70.12	42.36	94.42	81.90
$C_m^{(PO_\mu, PO_\mu)}$	95.97	85.62	40.58	83.98	95.18	93.66	97.52	89.72	83.57	71.25	41.54	94.21	81.07
$C_m^{(PO_\nu, PO_\nu)}$	96.64	85.43	43.62	82.70	95.20	93.97	97.43	99.03	83.75	69.36	42.31	94.20	81.97
$C_m^{(PO_\pi, PO_\pi)}$	95.97	85.43	42.31	82.65	95.23	93.48	97.53	99.21	84.85	70.24	41.85	94.35	81.93

Table 6. Ensemble algorithm performance of POP Choquet-like integrals that choose pairs with two different pseudo overlap functions.

Dataset Integral	IR	BS	WR	WF	OR	C9	WI	SG	CE	MH	WW	PB	Mean
$C_m^{(PO_\beta, PO_\alpha)}$	97.31	82.82	49.50	84.88	95.67	95.89	97.94	98.63	87.99	50.38	45.59	95.08	81.81
$C_m^{(PO_\beta, PO_\chi)}$	97.31	82.77	50.00	85.00	94.49	95.98	97.84	90.44	87.70	44.19	45.23	94.08	80.42
$C_m^{(PO_\beta, PO_\delta)}$	95.92	83.04	49.97	84.74	95.49	95.90	98.04	97.50	87.94	51.29	45.56	95.27	81.72
$C_m^{(PO_\beta, PO_\epsilon)}$	95.26	83.03	49.81	84.56	95.90	96.18	97.94	95.87	88.12	64.28	45.41	95.39	82.65
$C_m^{(PO_\beta, PO_\phi)}$	95.27	82.78	49.88	84.98	95.54	96.09	98.94	98.63	87.72	51.08	45.53	95.00	81.79
$C_m^{(PO_\chi, PO_\lambda)}$	95.97	86.34	46.94	83.70	96.71	94.98	97.59	98.34	86.85	65.48	41.89	95.25	82.50
$C_m^{(PO_\mu, PO_\phi)}$	95.97	86.35	44.67	84.38	96.74	93.85	97.64	99.37	87.04	72.51	42.27	94.84	82.97
$C_m^{(PO_\psi, PO_\nu)}$	95.97	86.35	41.48	83.75	96.47	93.56	97.64	98.90	86.33	62.81	40.71	94.75	81.56
$C_m^{(PO_\alpha, PO_\psi)}$	96.64	81.13	48.60	85.15	96.00	95.47	97.04	99.59	88.03	66.99	42.68	94.46	82.65
$C_m^{(PO_\iota, PO_\alpha)}$	93.23	60.78	50.50	81.54	92.48	95.55	98.97	90.46	85.04	35.86	43.67	95.08	76.93
$C_m^{(PO_\kappa, PO_\alpha)}$	91.81	60.54	51.60	81.06	92.67	95.83	98.04	91.54	85.34	35.67	46.16	95.14	77.12
$C_m^{(PO_\gamma, PO_\alpha)}$	97.31	82.97	49.17	85.36	95.67	95.61	84.04	98.63	87.84	66.02	45.34	95.31	81.94
$C_m^{(PO_\gamma, PO_\delta)}$	98.66	82.97	49.40	85.23	95.95	95.89	98.04	98.78	87.84	65.82	45.18	95.44	83.27
$C_m^{(PO_\gamma, PO_\eta)}$	96.64	82.70	49.02	85.34	96.04	95.80	98.04	99.25	87.94	66.42	45.47	94.86	83.13
$C_m^{(PO_\alpha, PO_\pi)}$	95.97	86.34	47.81	84.06	96.69	93.78	97.74	98.59	86.80	64.67	42.08	95.06	82.47

It can be seen that the results of ensemble classifiers based on pseudo-overlap function pairs are similar, which indicates the stability of this algorithm. Even if the pseudo overlap function pair with optimal performance is not selected, the final classification result is still acceptable. In particular, POP Choquet-like integrals that choose pairs with two different pseudo-overlap functions perform slightly better in the ensemble than POP Choquet-like integrals constructed with two identical pseudo-overlap functions. In Table 6, when the pseudo-overlap function pair (PO_γ, PO_δ) is selected, the average F1-Measure value of the final classification of the 12 data sets is the largest (83.27), and the classification effect is the best. In the following experiment, the experimental results of POP Choquet-like integral $C_m^{(PO_\gamma, PO_\delta)}$ are compared with those of other classification algorithms.

To illustrate the advantages of this algorithm in classification applications, Table 7 compares this algorithm with other advanced classifier ensemble algorithms in the literature: classifier ensemble algorithms based on C_m^O (CQO) [11], the generalized mixed function classifier (GM) [28], random forest trees (RT) [29], XGBoost (XGB) [30], META-DES (META) [31], lightGBM (LGBM) [32], randomized reference classifier (RRC) [33], and CatBoost (CA) [34]. Note that some of the comparative experiments for eight out of the twelve datasets we selected can be found in [11], and the details of the other comparative experiments will be introduced in the next paragraph.

Table 7. Comparison of the results of the ensemble algorithm.

Algorithm Dataset	$C_m^{(PO_\gamma, PO_\delta)}$	CQO	GM	RT	XGB	META	LGBM	RRC	CAT
IR	98.66	95.31	95.98	96.65	95.98	95.98	94.64	95.31	95.31
BS	82.97	61.08	55.87	45.56	51.61	55.21	48.92	56.00	56.14
WR	49.40	29.07	29.70	25.65	29.74	27.47	29.81	28.39	26.37
WF	85.23	84.93	84.00	81.29	84.86	77.53	84.62	76.93	85.28
OR	95.95	98.53	98.37	95.53	97.48	98.40	97.68	98.37	95.19
C9	95.89	95.08	93.88	90.42	90.82	93.81	84.34	94.19	87.62
WI	98.04	97.54	98.37	95.63	96.89	97.32	94.32	95.47	96.54
SG	98.78	94.84	94.87	91.31	95.37	94.71	95.46	94.53	95.15
CE	87.84	58.43	55.99	52.21	55.05	51.41	59.18	51.75	46.38
MH	65.82	35.99	32.45	30.68	33.57	31.47	37.87	31.29	29.84
WW	45.18	25.71	24.55	23.18	25.45	22.54	25.01	22.43	22.15
PB	95.44	95.21	95.33	94.89	94.54	95.42	95.36	94.07	93.52
Average	83.27	72.64	71.61	68.58	70.59	70.11	70.55	69.89	69.12
Win-loss	9–3	1–11	1–11	0–12	0–12	0–12	0–12	0–12	1–11
<i>p</i> -value	-	0.0013	0.014	0.006	0.010	0.009	0.006	0.007	0.008

For CQO, we chose the configuration in [11]: Naive Bayes, classical k-NN with Euclidean distances, multilayer perceptron, and CART with a proportion of 1:3:3:3, and we compare the results of the best performing overlap function and fuzzy measure in [11] with the results of our algorithm. For GM, we chose the configuration in [28]: k-NN, decision trees, MLP, Naive Bayes, SVM with a proportion of 3:3:2:1:1. This article is implemented with sklearn, the dataset situation determines the number of trees, and other parameters use the default values of the library for RT. For XGB, using the softmax loss function, the maximum depth of the tree is set between 3 and 10, and the number of trees is set between 100 and 500, depending on the size of the data sets. For META and RRC, we selected the same configuration as our algorithm. For LGBM, the maximum depth of the tree is set to 10, and the number of leaves is less than 2^{10} . For CA, the maximum number of iterations is set to 3000.

In Table 7, we use the F1 measure, score, and the *p* value of the *t*-test to compare our ensemble algorithm with other algorithms. The two values of the score (win-loss) indicate the number of data sets whose F1 measure averages are greater than and less than the F1 measure averages of other algorithms, respectively.

It can be seen from Table 7 that our proposed ensemble algorithm has a better classification effect on most data sets (BS, WR, CE, MH, WW) than others. In particular, some data sets can greatly improve classification accuracy by using our ensemble algorithm in the case of other ensemble algorithms with general classification effects. Because the classification qualities of our algorithms on these data sets are much greater than that of other ensemble algorithms, the results of the sample T -test with them are significantly different.

5. Community Network Detection Algorithm Based on POP Choquet-like Integrals

5.1. Modularity

As the focus of many scholars in recent years, community network has been widely studied. In previous explorations, scholars have noted that those communities and structures inherent in a social network are the main goals of understanding the network. A social network structure is often unclear in a real complex network. There are always overlapping parts between communities, which adds to the charm of the community network detection problem and is also why scholars are interested in this problem.

Based on the above questions, the authors of [22] pointed out that social networks fall into three categories: classical community networks (where there is no overlap at all), crisp community networks (where there is overlap and each node of the overlap can belong to more than one community), and fuzzy community networks (where each node belongs to each community to some extent).

When the three situations of network community are defined, the optimal division method should be found. Because the number of communities to be divided is unknown, the problem of community network division is a clustering problem. As for how to evaluate the quality of the community clustering algorithm, modularity is a commonly used measurement method. To find the optimal solution, we can judge the quality of the network community division according to the modularity value.

Modularity, proposed by Girvan and Newman in [35], is used to measure the classical network community division scheme. In this paper, it is represented by Q_{GN} and defined as

$$Q_{GN} = \frac{1}{2m} \sum_{i,j \in V} \left[A_{ij} - \frac{k_i k_j}{2m} \right] \delta(c_i c_j),$$

where $G = (V, E)$ is a given network, $C = \{c_1, \dots, c_r\}$ is a partition, m is the number of edges of the network, k_i represents the degree of node i , A_{ij} is the adjacency matrix of the network before partition, if node i and node j belong to the same community after partition, $\delta(c_i c_j) = 1$; otherwise, it is 0. Under the premise of an unknown number of community partitions, the Q_{GN} modularity is the most classic and most commonly used measure of community network partition schemes.

The value range of modularity is $[-0.5, 1]$, and within this interval, the larger the modularity value, the better the clustering effect will be. In particular, when the value of modularity is greater than 0.3, the superiority of the current algorithm can be explained.

Considering that Q_{GN} modularity is more suitable for classical network communities, the network community required to be measured has no overlap. However, in the real network, a node often belongs to several division areas simultaneously, so the original Q_{GN} modularity cannot accurately determine the division scheme of these fuzzy networks. Many scholars are committed to improving the Q_{GN} modularity to make it more suitable for the fuzzy network community.

In [21], Nepusz et al. defined crisp modularity in fuzzy scenarios, represented by Q_T in this paper. It is defined as follows:

$$Q_T = \frac{1}{2m} \sum_{i,j \in V} \left[A_{ij} - \frac{k_i k_j}{2m} \right] s_{ij}.$$

The Q_T modularity improves the classic Q_{GN} modularity. The author replaced $\delta(c_i c_j)$ with s_{ij} , which represents the sum of the product of membership degrees of node i and j belonging to the same community, $s_{ij} = \sum_{c=1, \dots, r} \mu_{C_c}(i) \mu_{C_c}(j)$, and satisfies $\forall i \in V, \sum_{c=1, \dots, r} \mu_{C_c}(i) = 1$. Finally, other symbols are consistent with the Q_{GN} modularity.

Although in [21], the author had extended the classical Q_{GN} modularity to the fuzzy scene, it also has some shortcomings. If there is a node membership degree $\sum_{c=1}^r \mu_{C_c}(i) > 1$ in the fuzzy network community after partition, then Q_T modularity is not applicable.

Considering the above shortcomings, Gomez had improved the Q_{GN} modularity in [22] to fully fit the fuzzy framework. In this paper, the modularity proposed by Gomez is represented by Q_D , which is defined as follows

$$Q_D = \frac{1}{2m} \sum_{i,j \in V} \left[A_{ij} - \frac{k_i k_j}{2m} \right] G_G \{ G_O(\mu_{C_c}(i), \mu_{C_c}(j)) | c \in C \},$$

where G_O is a two-dimensional overlapping function, G_G is an n -dimensional grouping function, and other symbols have the same meaning as Q_{GN} modularity.

Gomez, though, is mindful of the overlap of the web community and used overlap and group function with good performance to improve Q_{GN} modularity. However, in processing node membership, the group function continues to use the maximum value to realize, which is not a non-average mean, which is the shortcoming of the research.

Considering the above deficiency, this paper uses the non-average POP Choquet-like integral to improve the modularity to obtain new modularity, which is defined as follows:

$$\tilde{Q} = \frac{1}{2m} \sum_{i,j \in V} \left[A_{ij} - \frac{k_i k_j}{2m} \right] C_m^{(PO_1, PO_2)} \{ PO(\mu_{C_c}(i), \mu_{C_c}(j)) | c \in C \},$$

where PO is a two-dimensional pseudo overlap function, and $C_m^{(PO_1, PO_2)}$ are the Choquet-like POP integrals constructed in this paper's third section. The meaning of the remaining symbols is consistent with the modularity Q_{GN} .

Note that our proposed modularity is not just a solution to the fuzzy network community problem. When the network community is classic or clear, the modularity and Q_{GN} modularity has the same effect.

Example 4. In [20], the author designed a simple fuzzy graph network, as shown in Figure 2. In this network, nodes are naturally divided into three communities, with overlaps between the three communities. Nodes 4 and 8 belong to multiple communities at the same time.

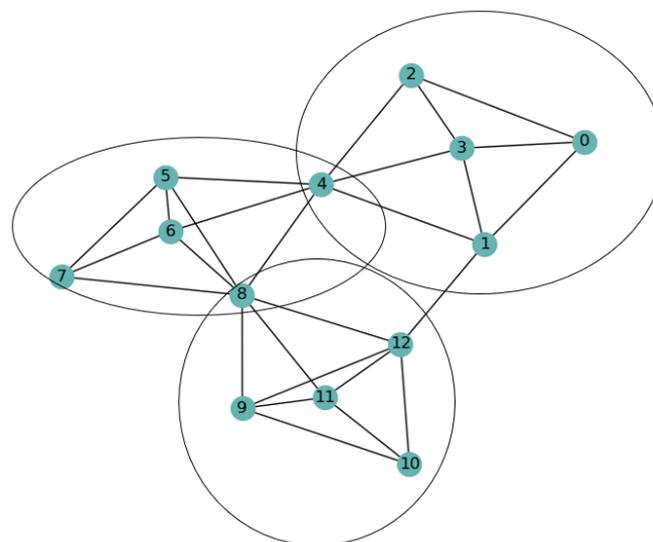


Figure 2. A simple fuzzy graph network for testing.

In [20], the author gives the degree of membership to each community after soft clustering of each node, as shown in Table 8. In this paper, the author stated that when the class number $C = 3$, the network modularity in the figure above was the highest (0.326).

Table 8. Soft clustering results of example graphs.

Nodes	Soft Clustering Results	Crisp C
0	[0.9951, 0.0026, 0.0023]	[1, 0, 0]
1	[0.9804, 0.0108, 0.0088]	[1, 0, 0]
2	[0.9984, 0.0008, 0.0008]	[1, 0, 0]
3	[0.9984, 0.0008, 0.0008]	[1, 0, 0]
4	[0.4327, 0.1133, 0.4540]	[1, 0, 1]
5	[0.0039, 0.0037, 0.9924]	[0, 0, 1]
6	[0.0039, 0.0037, 0.9924]	[0, 0, 1]
7	[0.0012, 0.0013, 0.9975]	[0, 0, 1]
8	[0.0715, 0.1519, 0.7766]	[0, 1, 1]
9	[0.0020, 0.9959, 0.0022]	[0, 1, 0]
10	[0.0012, 0.9976, 0.0011]	[0, 1, 0]
11	[0.0020, 0.9959, 0.0022]	[0, 1, 0]
12	[0.0054, 0.9899, 0.0047]	[0, 1, 0]

Bring the soft clustering results given in Table 8 into the modularity we defined for calculation, select the pseudo overlap function $\min\{x^{1/2}, y^{1/2}\}$, and randomly select the POP Choquet-like integral $C_m^{(PO_\beta, PO_\alpha)}$ (Using the g_λ fuzzy measure). Finally, when the partition number $C = 3$, the network modularity in the example figure is the highest (0.420), which can preliminarily verify that our modularity is reliable and advanced.

5.2. Experimental Framework

For several different network communities, many scholars have proposed different detection algorithms. However, few people have proposed non-average network community detection algorithms, so on the basis of [20,22], we propose a new community network detection algorithm based on the new modularity defined in Section 5.1. This process is explained in detail, and the overall algorithm framework is given below.

Algorithm 2: Pseudo-code for our proposed network community detection model

Input: An upper bound K and an adjacent matrix $A = (a_{ij})_{n \times n}$ for the number of clusters in a given network.

Output: The largest $\tilde{Q}(U_k)$ and its corresponding k .

- 1 **for** row in A **do**
- 2 $d_{row} = \text{sum}(\text{row})$
- 3 Generate a diagonal matrix D with diagonal element d : $D = \text{diag}(d)$;
- 4 Cholesky decomposition on G matrix: $G = \text{Cholesky}(d)$;
- 5 $a = G^{T(-1)}$;
- 6 $b = G^{(-1)} A G^{(-1)T}$;
- 7 **for** $k = 2 \rightarrow K$ **do**
- 8 Calculate the eigenvector of b : $e_1 = \text{eigvec}(b)$;
- 9 $E_k = a \cdot e_1$;
- 10 Make matrix $E_k = [e_2, \dots, e_k]$ from matrix $E_k = [e_1, \dots, e_k]$;
- 11 Use Euclidean distance norm is used to normalize the rows of E_k to unit lengths;
- 12 The soft distribution matrix U_k is obtained by clustering the row vectors of E_k with FCM;
- 13 Compute the $\tilde{Q}(U_k)$.
- 14 **return** The largest $\tilde{Q}(U_k)$ and its corresponding k and fuzzy partition U_k .

Fuzzy C-means (FCM) mentioned in the Algorithm 2 is a clustering method commonly used by scholars. It was proposed by Dunn as early as 1973. Later scholars tried to improve and put forward different FCM algorithms many times, especially the version proposed by Bezdek in [36], which has been used until now. FCM can allow each datapoint in the current cluster to belong to multiple classes simultaneously; a clustering algorithm fully adapted to the fuzzy framework. The main idea is to minimize the function

$$J_m = \sum_{i=1}^n \sum_{j=1}^k u_{ij}^m \|x_i - c_j\|^2,$$

where x_i is an n-dimensional data point to be clustered, and c_j is the n-dimensional clustering center of class j . u_{ij} is the degree to which x_i belongs to class j and satisfies $\sum_j u_{ij} = 1, m \in [1, \infty)$. $\|*\|$ represents any norm of similarity between any data point to be clustered and any clustering center.

Parameter m needs to be set for the FCM algorithm. An m value that is too large will lead to a poor clustering effect; an m value that is too small will make the algorithm similar to HCM and cannot highlight the fuzziness. In [37], the author proved that the best value range of m is [1.5,2.5]. When the only parameter to be set is confirmed, the FCM algorithm does not need human intervention in the implementation process.

In matrix theory [38], the generalized and ordinary eigenvalues of the same matrix are the same, and their eigenvectors are the same after normalization by Euclidean norm. However, it is a more stable numerical method to calculate the eigenvectors of generalized eigensystems. In the Algorithm 2 of this paper, the eigenvector of the $k-1$ dimensional generalized characteristic system of the diagonal matrix we calculated represents the $k-1$ dimensional values of the network graph, and these values serve as the numerical form of the points to be clustered.

5.3. Experimental Results and Analysis

To further test the benefits of our proposed social network detection algorithm, in this section, we compare it to other classic social network detection algorithms using two well-known reality networks: the Karate Club network and the Les Miserables network.

We use Python to implement the Algorithm 2 in this paper and select some classic network community detection algorithms and clear network community detection algorithms, such as GN [35] and D&L [39]. In addition, some advanced fuzzy network community detection algorithms are also used for comparisons, such as OCD [22] and NeSiFC [40].

In the process of algorithm implementation, we chose the pseudo overlap function $\min\{x^{1/2}, y^{1/2}\}$. Since there are too many POP Choquet-like integrals proposed in this paper, it would be too tedious to carry out experiments on all of them, so we chose four non-average POP Choquet-like integrals(Using the g_λ fuzzy measure) to carry out experiments, namely: $C_m^{(PO_\alpha, PO_\nu)}, C_m^{(PO_\alpha, PO_\varphi)}, C_m^{(PO_\varphi, PO_\nu)}, C_m^{(PO_\alpha, PO_\pi)}$.

(1) Network of Karate Club.

One of the most commonly used test networks for social network detection algorithms is the Karate Club network [41] (Figure 2), which Zachary observed over two years and has been mentioned in many articles. Karate Club is a real network with 34 members, each interacting with another member. The corresponding graph in the network is an undirected graph with 34 nodes and 78 edges. During these two years, because the relationship between the administrator and the coach broke down (nodes 0 and 34), the network was split into two smaller groups in reality, as shown in Figure 3.

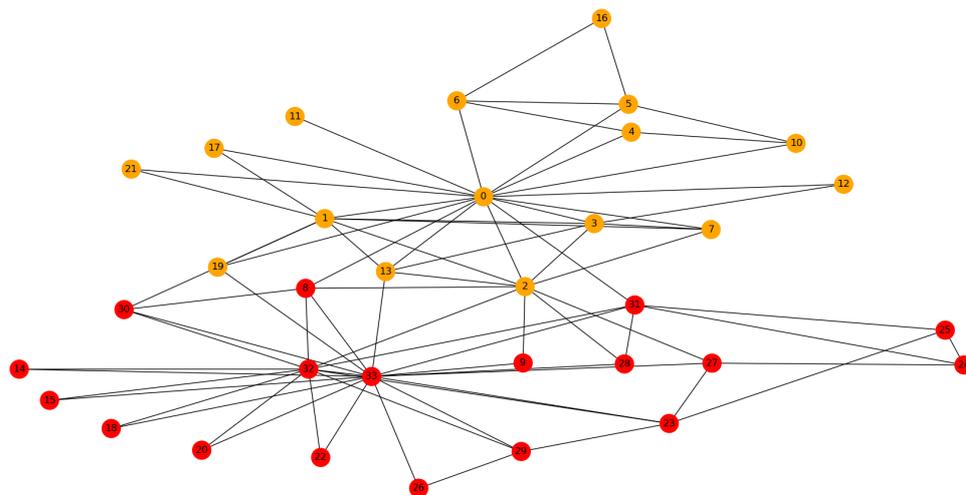


Figure 3. The Karate Club network.

It should be noted that social relationship networks will lead to fission because people’s emotions and transmitted information are not equal in society, which is constrained by interpersonal relationships and poor information, and other aspects. This inequality can lead to a network being divided into subgroups over time, with smaller subgroups being more stable. We want to identify potential information nodes where fission is likely to occur and reasonably predict what will happen next.

Table 9 shows the Karate Club network processing with the Algorithm 2 proposed in this paper, and the maximum modularity of each algorithm is shown in bold. The results show that the feedback obtained by the Algorithm 2 under the action of the four POP Choquet-like integrals is very close, and the differences are negligible. The measurement results of the Karate Club network show that when the partition number $C = 3$, the modularity value is the highest.

Table 9. Our algorithm measures the Karate Club network.

Choquet-like POP Integers Classes	$C_m^{(PO_\alpha, PO_\nu)}$	$C_m^{(PO_\alpha, PO_\varphi)}$	$C_m^{(PO_\varphi, PO_\nu)}$	$C_m^{(PO_\alpha, PO_\pi)}$
2	0.387	0.367	0.368	0.376
3	0.462	0.460	0.457	0.448
4	0.461	0.457	0.457	0.446
5	0.417	0.389	0.385	0.402
6	0.147	0.031	0.024	0.074
7	0.100	0.028	0.037	0.070
8	0.104	0.021	0.010	0.061
9	0.122	0.019	0.035	0.065
10	0.100	0.020	0.031	0.060
11	0.107	0.017	0.035	0.066
12	0.091	0.014	0.034	0.056
13	0.073	0.012	0.029	0.051
14	0.051	0.010	0.014	0.048
15	0.044	0.007	0.012	0.026
16	0.042	0.004	0.006	0.008
17	0.007	0.001	0.002	0.004

To illustrate the practicability of the proposed algorithm, Table 10 compares the best results obtained in Table 9 with other advanced algorithms. For the Karate Club network, the GN algorithm has the best partition effect when the partition number $C = 5$ (modularity value is 0.385). D&L and OCD algorithms perform best when the partition number $C = 4$ (modularity value: 0.416, 0.437). NeSiFC algorithm is a recently proposed network community detection algorithm based on neighbor similarity. It does not need to determine the partition number. In [40], the author used it to calculate the Karate Club network had a maximum modularity value of 0.372. As you can see, our network algorithm obtains a higher modularity value than other algorithms.

Table 10. Karate Club network experimental comparison results.

Classes	GN	D&L	OCD	NeSiFC	$C_m^{(PO_a, PO_v)}$
2	0.360	0.315	0.340		0.387
3	0.349	0.385	0.400		0.462
4	0.363	0.416	0.437		0.461
5	0.385	0.413	0.434		0.417
6	0.352	0.406	0.405		0.147
7	0.376	0.398	0.310	Best: 0.372	0.100
8	0.358	0.389	0.215		0.104
9	0.342	0.377	0.213		0.122
10	0.325	0.362	0.318		0.100
11	0.316	0.351	0.230		0.107
12	0.299	0.334	0.120		0.091
13	0.280	0.317	0.221		0.073
14	0.263	0.300	0.251		0.051
15	0.248	0.282	0.346		0.044
16	0.227	0.252	0.208		0.042
17	0.209	0.231	0.172		0.007

Figure 4 shows the node division of the algorithm in this paper. We use yellow, blue, and green to represent three different partitions to see which community each node belongs to. The criteria are that a node belongs to the community if its membership to the current community is greater than 0.25. The nodes in red represent overlapping nodes with membership greater than 0.25 to several communities simultaneously. Under the current partition, all three overlapping nodes belong to both yellow and green communities.

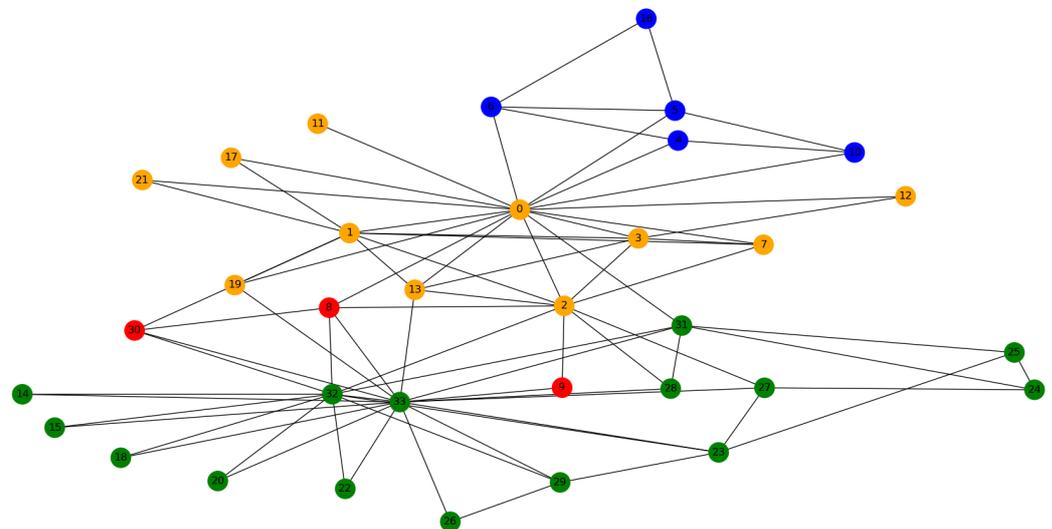


Figure 4. The clustering result of the Karate Club network.

(2) The Les Miserables network

The Les Miserables network comes from the famous novelist Hugo's novel Les Miserables. The original data can be found at <http://www.personal.umich.edu/~mejn/netdata/> (accessed on 12 January 2023). The network has 77 nodes and 254 edges corresponding to the characters and their relationships in the novel (Figure 5).

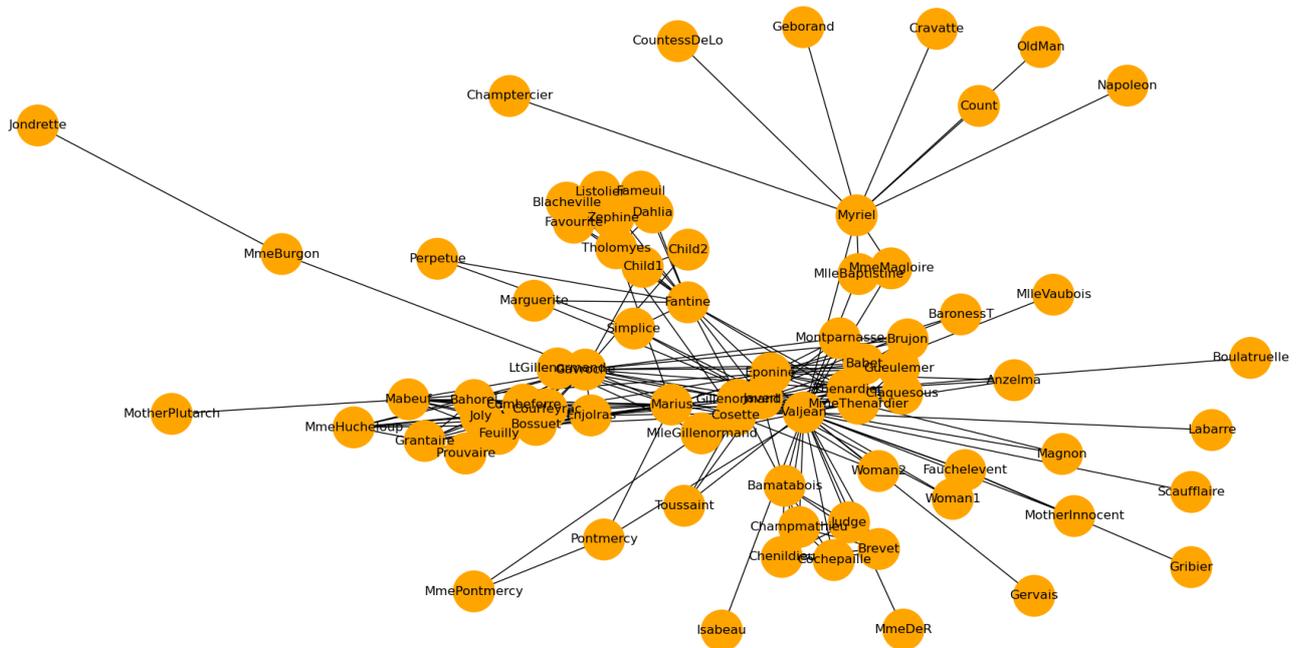


Figure 5. The Les Miserables network.

The degree of connection between wired nodes is the same as in the Karate Club network. That is, the values of its adjacency matrix are only 0 and 1. Unlike the Karate Club network, the Les Miserables network is much more complex and larger. That is, the value of the adjacency matrix of the Les Miserables network is not only 0 or 1 but also larger than 1.

Table 11 shows the measurement results of the Les Miserables network based on four different POP Choquet-like integrals in this paper. Similarly, it can be observed that the results of the four algorithms are similar, which shows the stability of our algorithm on the side. Table 11 shows that when the partition number $C = 5$, the modularity value is the highest, which means dividing five communities is the best fit for Les Miserables Network.

In Table 12, we compare the best results obtained in Table 11 with other advanced algorithms. Notice that for the Les Miserables network, the GN algorithm works best when $C = 11$ (modularity value is 0.538); D&L and OCD algorithms performed best when the partition number $C = 7$ (the modularity values were 0.556 and 0.564, respectively). In [40], the highest modularity value of the Les Miserables network calculated by the author using NeSiFC as 0.573. It can be seen that our proposed network algorithm obtains a higher modularity value (0.585) than other algorithms.

Table 11. Our algorithm measures the Les Miserables network.

Choquet-like POP Integers Classes	$C_m^{(PO_r, PO_v)}$	$C_m^{(PO_r, PO_\varphi)}$	$C_m^{(PO_\varphi, PO_v)}$	$C_m^{(PO_r, PO_\pi)}$
2	0.374	0.244	0.388	0.324
3	0.493	0.412	0.466	0.417
4	0.534	0.529	0.532	0.517
5	0.585	0.568	0.574	0.561
6	0.178	0.110	0.021	0.104
7	0.175	0.084	0.015	0.087
8	0.132	0.071	0.012	0.074
9	0.120	0.062	0.010	0.066
10	0.109	0.056	0.010	0.060
11	0.094	0.049	0.010	0.051
12	0.087	0.045	0.010	0.047
13	0.074	0.040	0.004	0.041
14	0.067	0.047	0.026	0.036
15	0.033	0.047	0.026	0.130
16	0.030	0.022	0.014	0.004
17	0.011	0.006	0.003	0.001

Table 12. Les Miserables Club network experimental comparison results.

Classes	GN	D&L	OCD	NeSiFC	$C_m^{(PO_r, PO_v)}$
2	0.075	0.372	0.233		0.374
3	0.260	0.464	0.264		0.493
4	0.267	0.511	0.494		0.534
5	0.416	0.552	0.553		0.585
6	0.459	0.554	0.556		0.178
7	0.456	0.556	0.564		0.175
8	0.454	0.556	0.276	Best: 0.573	0.132
9	0.452	0.553	0.260		0.120
10	0.452	0.551	0.113		0.109
11	0.538	0.548	0.233		0.094
12	0.535	0.546	0.174		0.087
13	0.531	0.543	0.115		0.074
14	0.528	0.540	0.061		0.067
15	0.525	0.537	0.041		0.033
16	0.523	0.525	0.026		0.030
17	0.520	0.520	0.041		0.011

Figure 6 shows the partition of the Les Miserables Network by our proposed network algorithm. The five communities are shown in pink, purple, blue, green, and yellow. Similarly, the threshold of 0.25 is used to determine whether the node belongs to the current community. The red nodes represent overlapping nodes. Node Marguerite belongs to both purple and yellow communities. Node Perpetue belongs to both the pink and yellow communities.

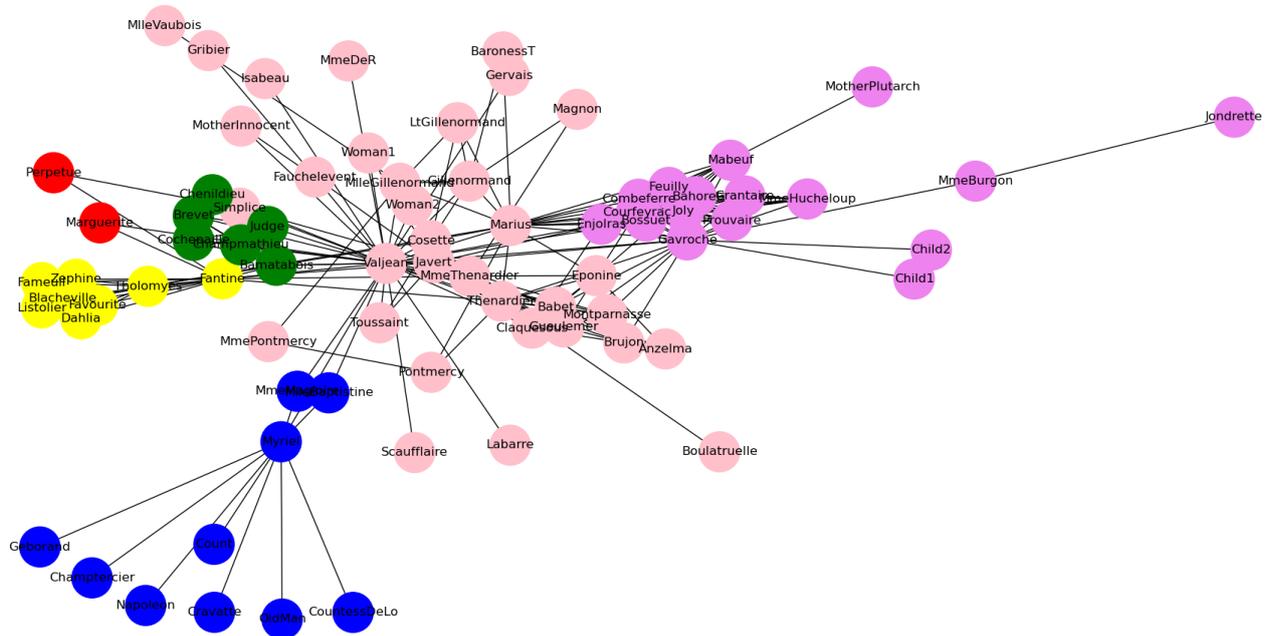


Figure 6. Clustering result of Les Miserables network.

6. Conclusions

In this paper, we proposed the POP Choquet-like integrals and applied them to classifier ensemble and fuzzy community detection, which makes good results. The contributions of this paper are listed as follows:

- The pseudo overlap function pair is introduced to replace products in discrete Choquet integral. So, the POP Choquet-like integral $C_m^{(PO_1, PO_2)}$ is obtained.
- Two new algorithms using the POP Choquet-like integral are designed. One is the ensemble algorithm, a branch of the classification algorithm. We use the defined $C_m^{(PO_1, PO_2)}$ as the fusion operator and the classification results of the base algorithms as inputs into the fusion operator to obtain a clear classification result. Another is the network community detection algorithm, a typical clustering algorithm. We use the defined $C_m^{(PO_1, PO_2)}$ to act on the results after each node's soft clustering, improving the classical modularity. Theoretically, our algorithm considers the non-average node membership degree in fuzzy community networks, which is more practical.
- Many experiments were conducted on multiple datasets, proving the advantages of the two algorithms.

In future research, the following topics deserve attention: (1) The POP Choquet-like integrals will be combined with other decision making and aggregation methods, such as neural networks, decision trees, etc. (2) The integration of the research approach of this paper with the latest research findings (such as [42–45]).

Author Contributions: Conceptualization, methodology, designing computer programs, X.Z.; writing—original draft preparation, designing computer programs, H.J.; writing—reviewing and editing, validation, J.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work is funded by the National Natural Science Foundation of China (Nos. 12271319 and 12201373), and China Postdoctoral Science Foundation 2023T160402

Data Availability Statement: All the data we used has been presented in the paper Table 4, Figures 2, 3 and 5.

Conflicts of Interest: No conflict of interest exists in the submission of this manuscript, and all authors approve the manuscript for publication. The work was original research that has not been published previously and is not under consideration for publication elsewhere, in whole or in part. All the authors listed have approved the manuscript that is enclosed.

References

1. A dankon, M.; Cheriet, M.; Biem, A. Semisupervised learning using Bayesian interpretation: Application to LS-SVM. *IEEE Trans. Neural Netw.* **2011**, *22*, 513–524. [[CrossRef](#)]
2. Vapnik, V. *The Nature of Statistical Learning Theory*; Springer Science & Business Media: New York, NY, USA, 1999.
3. Swaney, D.; McAlister G.; Coon J. Decision tree-driven tandem mass spectrometry for shotgun proteomics. *Nat. Methods* **2008**, *5*, 959–964. [[CrossRef](#)]
4. Valova, I.; Harris, C.; Mai, T. Optimization of convolutional neural networks for imbalanced set classification. *Procedia Comput. Sci.* **2020**, *176*, 660–669. [[CrossRef](#)]
5. Denoeux, T. A k-nearest neighbor classification rule based on Dempster-Shafer theory. *IEEE Trans. Syst. Man Cybern.* **1995**, *25*, 804–813. [[CrossRef](#)]
6. Tarus, J.; Ogada, K.; Mwalili, T. An Ensemble Based Model for Detecting Genetically Inherited Disorder. In Proceedings of the 2022 12th International Conference on Advanced Computer Information Technologies (ACIT), Ruzomberok, Slovakia, 26–28 September 2022; IEEE: New York, NY, USA, 2022; pp. 121–125.
7. AminiMotlagh, M.; Shahhoseini, H.; Fatehi, N. A reliable sentiment analysis for classification of tweets in social networks. *Soc. Netw. Anal. Min.* **2023**, *13*, 1–11. [[CrossRef](#)]
8. Qiu, H.; Su, P.; Jiang, S. Learning from Human Uncertainty by Choquet Integral for Optic Disc Segmentation. In Proceedings of the 2021 4th International Conference on Control and Computer Vision, Macau, China, 13–15 August 2021; pp. 7–12.
9. Tripathi, D.; Edla, D.; Cheruku, R. A novel hybrid credit scoring model based on ensemble feature selection and multilayer ensemble classification. *Comput. Intell.* **2019**, *35*, 371–394. [[CrossRef](#)]
10. Taha, A.; Barukab, O.; Malebary, S. Fuzzy Integral-Based Multi-Classifiers Ensemble for Android Malware Classification. *Mathematics* **2021**, *9*, 2880. [[CrossRef](#)]
11. Batista, T.; Bedregal, B.; Moraes, R. Constructing multi-layer classifier ensembles using the Choquet integral based on overlap and quasi-overlap functions. *Neurocomputing* **2022**, *500*, 413–421. [[CrossRef](#)]
12. Jain, A.; Murty, M.; Flynn, P. Data clustering: A review. *ACM Comput. Surv.* **1999**, *31*, 264–323. [[CrossRef](#)]
13. Sharma, K.; Seal, A. Spectral embedded generalized mean based k-nearest neighbors clustering with s-distance. *Expert Syst. Appl.* **2021**, *169*, 114326. [[CrossRef](#)]
14. Zhang, H.; Li, H.; Chen, N. Novel fuzzy clustering algorithm with variable multi-pixel fitting spatial information for image segmentation. *Pattern Recognit.* **2022**, *121*, 108–201. [[CrossRef](#)]
15. Esfandian, N.; Hosseinpour, K. A clustering-based approach for features extraction in spectro-temporal domain using artificial neural network. *Int. J. Eng.* **2021**, *34*, 452–457.
16. Liu, M.; Qi, X.; Pan, H. Construction of network topology and geographical vulnerability for telecommunication network. *Comput. Netw.* **2022**, *205*, 108764. [[CrossRef](#)]
17. Ouyang, X.; Huo, J.; Xia, L. Dual-sampling attention network for diagnosis of COVID-19 from community acquired pneumonia. *IEEE Trans. Med. Imaging* **2020**, *39*, 2595–2605. [[CrossRef](#)] [[PubMed](#)]
18. Liu, N.; Hu, X.; Ma, L. Vulnerability assessment for coupled network consisting of power grid and EV traffic network. *IEEE Trans. Smart Grid* **2021**, *13*, 589–598. [[CrossRef](#)]
19. Liu, X.; Maiorino, E.; Halu, A. Robustness and lethality in multilayer biological molecular networks. *Nat. Commun.* **2020**, *11*, 6043. [[CrossRef](#)]
20. Zhang, S.; Wang, R.; Zhang, X. Identification of overlapping community structure in complex networks using fuzzy c-means clustering. *Phys. A Stat. Mech. Appl.* **2007**, *374*, 483–490. [[CrossRef](#)]
21. Nepusz, T.; Petróczy, A.; Négyessy, L. Fuzzy communities and the concept of bridgeness in complex networks. *Phys. Rev. E* **2008**, *77*, 016107. [[CrossRef](#)]
22. Gomez, D.; Rodríguez, J.; Yanez, J. A new modularity measure for fuzzy community detection problems based on overlap and grouping functions. *Int. J. Approx. Reason.* **2016**, *74*, 88–107. [[CrossRef](#)]
23. Lucca, G.; Dimuro, G.; Fernández, J. Improving the Performance of Fuzzy Rule-Based Classification Systems Based on a Nonaveraging Generalization of CC-Integrals Named $C_{F_1 F_2}$ -Integrals. *IEEE Trans. Fuzzy Syst.* **2018**, *27*, 124–134. [[CrossRef](#)]
24. Zhang, X.; Liang, R.; Bustince, H. Pseudo overlap functions, fuzzy implications and pseudo grouping functions with applications. *Axioms* **2022**, *11*, 593. [[CrossRef](#)]
25. Choquet, G. Theory of capacities. *Ann. Inst. Fourier* **1954**, *5*, 131–295. [[CrossRef](#)]
26. Li, X.; Wang, F.; Chen, X. Support vector machine ensemble based on Choquet integral for financial distress prediction. *Int. J. Pattern Recognit. Artif. Intell.* **2015**, *29*, 1550016. [[CrossRef](#)]

27. Qiao, J.; Hu, B. On homogeneous, quasi-homogeneous and pseudo-homogeneous overlap and grouping functions. *Fuzzy Sets Syst.* **2019**, *357*, 58–90. [[CrossRef](#)]
28. Costa, V.; Farias, A.; Bedregal, B. Combining multiple algorithms in classifier ensembles using generalized mixture functions. *Neurocomputing* **2018**, *313*, 402–414. [[CrossRef](#)]
29. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
30. Chen, T.; Guestrin, C. Xgboost: A scalable tree boosting system. In Proceedings of the 22nd ACM Sigkdd International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.
31. Cruz, R.; Sabourin, R.; Cavalcanti, G. META-DES: A dynamic ensemble selection framework using meta-learning. *Pattern Recognit.* **2015**, *48*, 1925–1935. [[CrossRef](#)]
32. Ke, G.; Meng, Q.; Finley, T. Lightgbm: A highly efficient gradient boosting decision tree. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 3146–3154.
33. Woloszynski, T.; Kurzynski, M. A probabilistic model of classifier competence for dynamic ensemble selection. *Pattern Recognit.* **2011**, *44*, 2656–2668. [[CrossRef](#)]
34. Prokhorenkova, L.; Gusev, G.; Vorobev, A. CatBoost: Unbiased boosting with categorical features. *Adv. Neural Inf. Process. Syst.* **2018**, *31*, 6639–6649.
35. Girvan, M.; Newman, M. Community structure in social and biological networks. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 7821–7826. [[CrossRef](#)]
36. Bezdek, J.; Ehrlich, R.; Full, W. FCM: The fuzzy c-means clustering algorithm. *Comput. Geosci.* **1984**, *10*, 191–203. [[CrossRef](#)]
37. Pal, N.; Bezdek, J. On cluster validity for the fuzzy c-means model. *IEEE Trans. Fuzzy Syst.* **1995**, *3*, 370–379. [[CrossRef](#)]
38. Franklin, J. *Matrix Theory*; Courier Corporation: North Chelmsford, MA, USA, 2012.
39. Gomez, D.; Zarrazola, E.; Yanez, J. A divide-and-link algorithm for hierarchical clustering in networks. *Inf. Sci.* **2015**, *316*, 308–328. [[CrossRef](#)]
40. Roy, U.; Muhuri, P.; Biswas, S. NeSiFC: Neighbors' similarity-based fuzzy community detection using modified local random walk. *IEEE Trans. Cybern.* **2021**, *52*, 10014–10026. [[CrossRef](#)]
41. Zachary, W. An information flow model for conflict and fission in small groups. *J. Anthropol. Res.* **1977**, *33*, 452–473. [[CrossRef](#)]
42. Wang, J.; Zhang, X. A novel multi-criteria decision-making method based on rough sets and fuzzy measures. *Axioms* **2022**, *11*, 275. [[CrossRef](#)]
43. Zhang, X.; Sheng, N.; Borzooei, R.A. Partial residuated implications induced by partial triangular norms and partial residuated lattices. *Axioms* **2023**, *12*, 63. [[CrossRef](#)]
44. Zhang, X.; Li, M.; Liu, H. Overlap Functions-Based Fuzzy Mathematical Morphological Operators and Their Applications in Image Edge Extraction. *Fractal Fract.* **2023**, *7*, 465. [[CrossRef](#)]
45. Jing, M.; Zhang, X. Pseudo-Quasi Overlap Functions and Related Fuzzy Inference Methods. *Axioms* **2023**, *12*, 217. [[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.