*mathematics*

MDPI

*Article*

# Towards Dynamic Fuzzy Rule Interpolation via Density-Based Rule Promotion from Interpolated Outcomes

**Jinle Lin, Changjing Shang * and Qiang Shen**

Department of Computer Science, Aberystwyth University, Aberystwyth SY23 3DB, UK; jil29@aber.ac.uk (J.L.); qqs@aber.ac.uk (Q.S.)
* Correspondence: cns@aber.ac.uk

**Abstract:** Traditional fuzzy rule-based systems struggle with scenarios where knowledge gaps exist in the problem domain, due to limited data or experience. Fuzzy rule interpolation (FRI) effectively addresses the challenge of inference in fuzzy systems when faced with unmatched observations, due to the employment of an incomplete or sparse rule base. It generates temporary, interpolated rules for the unmatched observations, ensuring continued inference capability. However, the resultant valuable interpolated rules are conventionally discarded. This paper introduces a formal approach for dynamic fuzzy rule interpolation (D-FRI), based on the concept of density-based rule promotion and assisted by the use of the OPTICS clustering algorithm, through exploiting frequently appearing interpolated rules on the fly. This enhances the system's knowledge coverage, efficiency, and robustness over time. An implementation of such a D-FRI system is presented, which combines transformation-based fuzzy rule interpolation (T-FRI) with OPTICS clustering. This offers an effective mechanism for evaluating and subsequently selecting potentially powerful interpolated rules for the system to dynamically enrich its knowledge base. The implemented system is verified by experimental investigations.

**Keywords:** fuzzy rule interpolation; dynamic rule interpolation; OPTICS clustering; density-based rule promotion

**MSC:** 37M99

check for updates

**Citation:** Lin, J.; Shang, C.; Shen, Q. Towards Dynamic Fuzzy Rule Interpolation via Density-Based Rule Promotion from Interpolated Outcomes. *Mathematics* **2024**, *12*, 402. https://doi.org/10.3390/math12030402

Academic Editor: Hsien-Chung Wu

Received: 18 December 2023
Revised: 20 January 2024
Accepted: 24 January 2024
Published: 26 January 2024

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

## 1. Introduction

In a traditional fuzzy inference system with a sparse rule base, the absence of a corresponding rule for an observation results in the failure to generate any inferred outcome. A sparse rule base means the inference system lacks certain knowledge coverage of a problem domain. Furthermore, when applied to real-world problems, a static rule base may also lose its relevance and effectiveness due to the nature of the dynamics of the underlying changing world. Fuzzy rule interpolation (FRI) approaches introduce a novel way to enhance the reasoning potential when the rule base is incomplete, especially when it can dynamically update itself. This significantly supports type-1 fuzzy systems that deal with simple membership functions, involving less computational complexity while requiring less domain-specific knowledge [1–5].

In general, there exist two categories of fuzzy rule interpolation (FRI) methodologies. The first approach is centred on the utilisation of $\alpha$-cuts, whereby the $\alpha$-cut of an interpolated consequent is calculated from the $\alpha$-cuts of its antecedents, as outlined by the resolution principle [6–9]. The second methodology employs the concept of similarity of fuzzy value and analogical reasoning [10–17]. Specifically, within this latter category, an FRI algorithm generates intermediate fuzzy rules based on the principle of similarity. Among such methods, transformation-based fuzzy rule interpolation (T-FRI) is the most popular [12]. T-FRI can manage multiple fuzzy rules in both rule interpolation and extrapolation while guaranteeing the uniqueness, convexity, and normality of the interpolated fuzzy

sets. Furthermore, it accommodates multiple antecedent variables and varying shapes of membership functions. Consequently, this paper takes classical T-FRI techniques to generate interpolated rules, forming the foundation for a novel dynamic FRI methodology.

In the seminal T-FRI method, intermediate rules artificially generated through fuzzy rule interpolation (FRI) are discarded once an interpolated outcome is derived. Nonetheless, these interpolated rules may embed information that holds the potential to extend the coverage of a sparse rule base. Therefore, developing a methodology for collecting the interpolated rules generated during the FRI procedure is imperative for formulating a dynamic fuzzy rule interpolation mechanism.

In dynamic fuzzy reasoning, adaptive rule-based methodologies have gained popularity for traditional fuzzy systems that employ a dense rule base. These systems often apply various optimisation techniques to improve inference accuracy by cultivating a dynamic and dense rule base [18–23]. However, these adaptive strategies, built upon a dense rule base, can hardly apply to sparse rule-based fuzzy systems where the rules are inadequate for comprehensive coverage of the problem domain. Unfortunately, approaches designed for sparse rule-based dynamic fuzzy rule interpolation are rather limited. One such approach is fundamentally framed within a predefined quantity space of domain variables, called hypercubes [24]. This technique employs K-means clustering and genetic algorithms (GAs) to sort through hypercubes of interpolated rules. Subsequently, it generates a new rule from the selected hypercube via a weighting process. Whilst simple in conception, it forms the foundation of dynamic fuzzy rule interpolation (D-FRI) strategies and is fundamental for future methodologies that seek to manipulate pools of interpolated rules.

Learning from interpolated rules involving multiple antecedent variables can be interpreted as a complex task of optimisation or clustering. Techniques like genetic algorithms (GAs) and particle swarm optimisation (PSO) are prominent in the domain of systems learning and optimisation. Meanwhile, clustering methodologies such as K-means, mean-shift, density-based spatial clustering of applications with noise (DBSCAN) [25,26], and ordering points to identify the clustering structure (OPTICS) [27] exist. Notably, OPTICS is an improved version of DBSCAN as a density-based clustering algorithm. Unlike K-means, where every data point is ultimately categorised into a cluster, OPTICS can filter out noisy data, effectively excluding noise from the constructed clusters. When employed to select rules from a pool of interpolated rules, OPTICS can help mitigate the influence of outliers and assemble clusters largely populated by commonly encountered observations. Consequently, the interpolated results within these clusters are more generalised and reliable for referencing in a dynamic system. Inspired by this observation, the present work applies OPTICS to implement a novel approach to D-FRI.

Dynamic fuzzy rule interpolation is particularly effective in domains where knowledge is either incomplete or sparse, and the environment is either dynamic or non-stationary. Examples include real-time decision-making, adaptive control, and problems requiring robust solutions with limited or evolving data, such as financial or intelligence data analysis [24], and environmental or medical risk monitoring [28]. The main objective of dynamic fuzzy rule interpolation (D-FRI), in the context of this paper, is to enhance the inference capability of fuzzy rule-based systems, especially in situations where the rule base incompletely covers the problem domain. Interpolated outcomes are temporary rules generated during the process of running classical fuzzy rule interpolation when an observation does not match any existing rules in the rule base. Such outcomes, therefore, fill the existing knowledge gaps and allow the overall fuzzy system to make inferences when encountering similar observations in the future. Particularly in D-FRI, the interpolated outcomes are evaluated, with those of a sufficient occurrence frequency being promoted into the rule base, thus enhancing its dynamic coverage of the underlying domain. In so doing, the system's knowledge coverage, efficiency, and robustness are improved over time.

The key contributions of this work are multi-fold: (1) A novel framework for dynamic fuzzy rule interpolation. (2) An innovative integration of T-FRI and OPTICS clustering. (3) A density-based rule promotion procedure that enables the system to identify and

promote the most frequently occurring and relevant interpolated rules, thus bridging the gap between static rule-based systems and the need for adaptive inference mechanisms in changing environments, thereby advancing the practical applicability of fuzzy rule-based systems. (4) An initial implementation and experimental study of the integrated system that demonstrates the potential of developing practically feasible fuzzy rule-based systems with limited knowledge.

The paper is structured as follows. Section 2 introduces the core computational mechanisms of T-FRI and those of OPTICS. Section 3 illustrates the proposed OPTICS-assisted density-based dynamic fuzzy rule interpolation framework. Section 5 provides experimental results over different datasets. Section 5.3 presents a discussion regarding the experimental results. Finally, Section 6 concludes the paper and outlines the potential improvements that would be interesting to make in the future.

## 2. Foundational Work

This section provides an overview of the fundamental principles and procedural steps of T-FRI and OPTICS. Within the context of this paper, as with many recent techniques developed in the literature, triangular membership functions are employed to represent fuzzy sets, both for popularity and simplicity.

### *2.1. Transformation-Based Fuzzy Rule Interpolation (T-FRI)*

Without losing generality, the structure of a sparse rule base can be denoted as $\mathcal{R}$, which includes $R_i \in \mathcal{R}$ fuzzy rules [12]. The specific rule format is delineated as follows:

$$R_i : \text{IF } x_1 \text{ is } A_{i,1} \text{ AND ... AND } x_N \text{ is } A_{i,N} \text{ THEN } y \text{ is } B_i,$$

where $i$ spans the set $\{1, \ldots, |\mathcal{R}|\}$, with the term $|\mathcal{R}|$ signifying the total count of rules present in the rule base. In the given domain, each $x_j$ with $j \in \{1, \ldots, N\}$ represents an antecedent variable, with $N$ indicating the overall number of these variables. The notation $A_{i,j}$ indicates the linguistic value attributed to the variable $x_j$ within the rule $R_i$, and it can be represented as the triangular fuzzy set $(a_{0,j}, a_{1,j}, a_{2,j})$. Here, $a_{0,j}$ and $a_{2,j}$ define the left and right boundaries of the support (where the membership values become zero), and $a_{1,j}$ denotes the peak of the triangular fuzzy set (with a membership value of one).

Considering an instance $O$, it is typically denoted by the sequence $A_{O,1}, \ldots, A_{O,j}$ with $j \in \{1, \ldots, N\}$. In this sequence, $A_{O,j} = (a_0, a_1, a_2)$ represents a specific triangular fuzzy value of the variable $x_j$. For any triangular fuzzy set $A$, defined as $(a_0, a_1, a_2)$, its representative value, termed as $rep(A)$, can be described as the average of its three defining points, which is given by:

$$rep(A) = (a_0 + a_1 + a_2)/3 \tag{1}$$

Building upon the foundational concepts outlined above, T-FRI employs the following key procedures when observation fails to match any rules within the sparse rule base.

**Closest rule determination**. The Euclidean distance metric is applied while measuring distances between a rule and an observation (or between two rules). The algorithm selects the rules closest to the unmatched observation to initiate the interpolation procedure. The distance between a rule $R_i$ and the observation $O$ is computed by aggregating the component distances between them, as shown below:

$$d(R_i, O) = \sqrt{\sum_{j=1}^{N} d_j^2}, \quad d_j = \frac{d(A_{i,j}, A_{O,j})}{range_{x_j}} \tag{2}$$

where $d(A_{i,j}, A_{O,j}) = |rep(A_{i,j}) - rep(A_{O,j})|$ defines the distance between the representative values in the domain of the $j$th antecedent variable. Additionally, the term $range_{x_j}$ is calculated as the difference between the maximum and minimum values of the variable $x_j$ within its domain.

**Intermediate rule construction**. The normalised displacement factor, symbolised as $\omega_{i,j}$, represents the significance or weight of the $j$th antecedent in the $i$th rule:

$$\omega_{i,j} = \frac{\omega_{i,j}^{\dagger}}{\sum_{i=1}^{M} \omega_{i,j}^{\dagger}} \tag{3}$$

where $M$ (with $M \geq 2$) indicates the count of rules selected based on their minimum distance values related to observation $O$. The weight $\omega_{i,j}^{\dagger}$ is given by:

$$\omega_{i,j}^{\dagger} = \frac{1}{1 + d(A_{O,j}, A_{i,j})} \tag{4}$$

Intermediate fuzzy terms, denoted as $A_j^{\dagger\dagger}$, are calculated utilising the antecedents from the nearest $M$ rules:

$$A_j^{\dagger\dagger} = \sum_{i=1}^{M} \omega_{i,j} A_{i,j} \tag{5}$$

These intermediate fuzzy terms are subsequently adjusted to $A_j^{\dagger}$ to ensure their representative values aligning with those of $A_{O,j}$:

$$A_j^{\dagger} = A_j^{\dagger\dagger} + \delta_j range_{x_j} \tag{6}$$

where $\delta_j$ quantifies the offset between $A_{O,j}$ and $A_j^{\dagger}$ in the domain of the $j$th variable:

$$\delta_j = \frac{rep(A_{O,j}) - rep(A_j^{\dagger})}{range_{x_j}} \tag{7}$$

From this, the intermediate consequent, symbolised as $B^{\dagger}$, is derived similarly to Equation (6). The aggregated weights $\omega_{B_i}$ and shift $\delta_B$ are calculated from the respective values of $A_j^{\dagger}$, such that

$$\omega_{B_i} = \frac{1}{N} \sum_{j=1}^{N} \omega_{i,j}, \quad \delta_B = \frac{1}{N} \sum_{j=1}^{N} \delta_j \tag{8}$$

**Scale transformation**. Let $A_{Sj}$ be represented by the fuzzy set $(a_0^{\dagger\dagger}, a_1^{\dagger\dagger}, a_2^{\dagger\dagger})$, resulting from a scale transformation applied to the $j$th intermediate antecedent value $A_j^{\dagger} = (a_0^{\dagger}, a_1^{\dagger}, a_2^{\dagger})$. Given a scaling ratio $s_j$, the components of $A_{Sj}$ are defined as:

$$a_0^{\dagger\dagger} = \frac{a_0^{\dagger}(1 + 2s_j) + a_1^{\dagger}(1 - s_j) + a_2^{\dagger}(1 - s_j)}{3} \tag{9}$$

$$a_1^{\dagger\dagger} = \frac{a_0^{\dagger}(1 - s_j) + a_1^{\dagger}(1 + 2s_j) + a_2^{\dagger}(1 - s_j)}{3} \tag{10}$$

$$a_2^{\dagger\dagger} = \frac{a_0^{\dagger}(1 - s_j) + a_1^{\dagger}(1 - s_j) + a_2^{\dagger}(1 + 2s_j)}{3} \tag{11}$$

This means that the scale rate $s_j$ is determined by

$$s_j = \frac{a_2^{\dagger\dagger} - a_0^{\dagger\dagger}}{a_2^{\dagger} - a_0^{\dagger}} \tag{12}$$

For the corresponding consequent, the scaling factor $s_B$ is then computed using

$$s_B = \frac{\sum_{j=1}^{N} s_j}{N} \tag{13}$$

**Move transformation**. $A_{Sj}$, after performing a scale transformation, is then shifted based on a move rate $m_j$. This adjustment ensures that the final scaled and shifted fuzzy set retains the geometric properties of the observed value $A_{O,j}$. In implementation, mathematically, this implies that the move rate $m_j$ is determined such that

$$m_j = \frac{3(a_0 - a_0^{\dagger\dagger})}{a_1^{\dagger\dagger} - a_0^{\dagger\dagger}} \tag{14}$$

where $a_0$ is the lower limit of the support of $A_{O,j} = (a_0, a_1, a_2)$. From this, the move factor $m_B$ for the consequent can then be derived as follows:

$$m_B = \frac{\sum_{j=1}^{N} m_j}{N} \tag{15}$$

**Interpolated Consequent Calculation**. In adherence to the analogical reasoning principle, the determined parameters $s_B$ (from the scaling transformation) and $m_B$ (from the move transformation) are collectively utilised to calculate the consequent of the intermediary rule $B^{\dagger}$. As a result, the interpolated outcome $B_O$ is generated in reaction to the unmatched observation in question. This methodology ensures that the computed output by the FRI algorithm, while derived from intermediate rules, closely reflects the characteristics of the given observation, thus maintaining the intuition behind the entire inference process.

### *2.2. OPTICS Clustering*

OPTICS, a data density-based clustering algorithm [27], builds upon the foundation of DBSCAN [25], aiming to identify clusters of varying densities. Notable advantages of OPTICS over K-means and DBSCAN include its adeptness at detecting clusters of various densities, its independence from pre-specifying the number of clusters (a requirement in K-means), its ability to identify clusters of non-spherical shapes (unlike the sphere-bound K-means), and its proficiency in distinguishing between noise and clustered data. This last trait is shared with DBSCAN but contrasts with K-means, which assumes every data point belongs to one and only one cluster.

In addition, OPTICS provides a hierarchical cluster structure, allowing for nested clusters, a feature missing in both K-means and DBSCAN. This is measured by the parameter $\xi$, the minimum steepness on the reachability plot that forms a cluster boundary. However, OPTICS can be sensitive to its parameter settings, particularly the minimum number of points needed to define a cluster. Its computational complexity may also be more demanding because, unlike DBSCAN, where the maximum radius of the neighbourhood $\varepsilon$ is fixed, in OPTICS, $\varepsilon$ is devised to identify clusters across all scales, especially with large datasets. Nonetheless, in practice, for most datasets and specifically with an appropriate indexing mechanism, OPTICS runs efficiently, often close to $O(n \log n)$. Yet, it should be noted that degenerate cases can push the complexity closer to $O(n^2)$. For completion, Table 1 presents the basic concepts and parameters concerning OPTICS and Figure 1 provides a simple illustration of its working.

Formally, Algorithm 1 outlines the process of OPTICS. Also, as an illustration, Figure 2 shows the impact of different parameter settings upon the results of OPTICS clustering (as compared to that which employs DBSCAN) running on the classical Iris dataset [29]. Note that both algorithms share the same *MinSample* in defining the minimum size of a cluster; while DBSCAN relies purely on a fixed epsilon value, OPTICS uses $\xi$ to define reachability without such a fixed value. The reachability plot and the $\xi$ parameter are central to the OPTICS algorithm, providing a means to visualize and automatically identify clusters in data. In particular, the reachability plot shows the proximity of points in a space ordered by OPTICS, and the $\xi$ parameter helps to define what is considered a cluster based on the steepness of slopes in this plot.

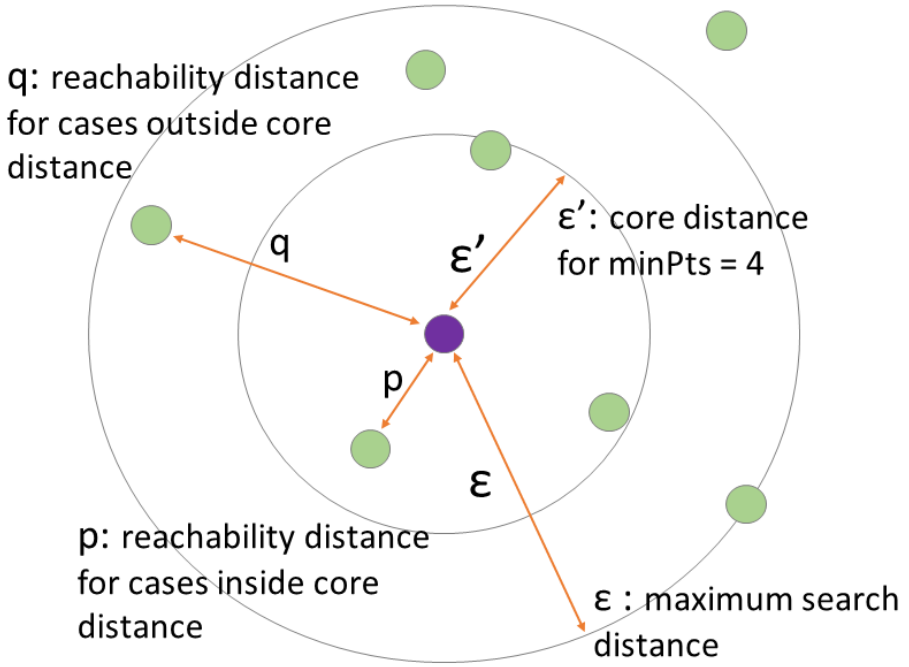



**Figure 1.** Illustration of OPTICS.

**Table 1.** Concepts and parameters of the OPTICS clustering algorithm.

| Concept | Definition |
|---|---|
| $\xi$ | Parameter used to specify the steepness that defines a cluster in the reachability plot, with clusters determined as regions in the reachability plot that exhibit a steep decrease, followed by a gentler increase in reachability distance. |
| MinPts | The minimum number of points needed to form a dense region, signifying that at least the MinPts are contained within its $\varepsilon$-neighbourhood for it to be considered a core point. |
| Reachability Distance | Smallest distance to reach a point while considering density (number of points) in the neighbourhood. |
| Core Distance | Smallest radius required for a point to cover the MinPts in its neighbourhood. |
| Ordering | Sequence in which points are processed, resulting in cluster ordering rather than explicit cluster assignments. |

**Algorithm 1** OPTICS clustering algorithm.

**Start**:
**Procedure OPTICS**(Dataset $D$, MinPts, $\xi$, maxSamples):
    **Step 1.** Initialise all points in $D$ as unprocessed.
        Set $OrderedList = \emptyset$ and $SeedList$ as an empty priority queue.
    **Step 2.** For each *point* in $D$, if *point* is unprocessed:
        Obtain neighbours with **RegionQuery** and process *point*.
        If *point* is a core point:
            Add its unprocessed neighbours to $SeedList$ with their respective reachability distances.
    **Step 3.** While $SeedList$ is not empty:
        Pop point $p$ with the smallest reachability distance from $SeedList$.
        Mark $p$ as processed and add to $OrderedList$.
        If $p$ is a core point:
            For each unprocessed neighbour $n$ of $p$:
                Update the reachability distance of $n$ and add/update its position in $SeedList$.
    **Step 4.** Return $OrderedList$.
**End.**
**Function RegionQuery**(*point*, $\xi$, maxSamples):
    Return all points within reachability distance considering $\xi$ and restricted by maxSamples.
**End.**

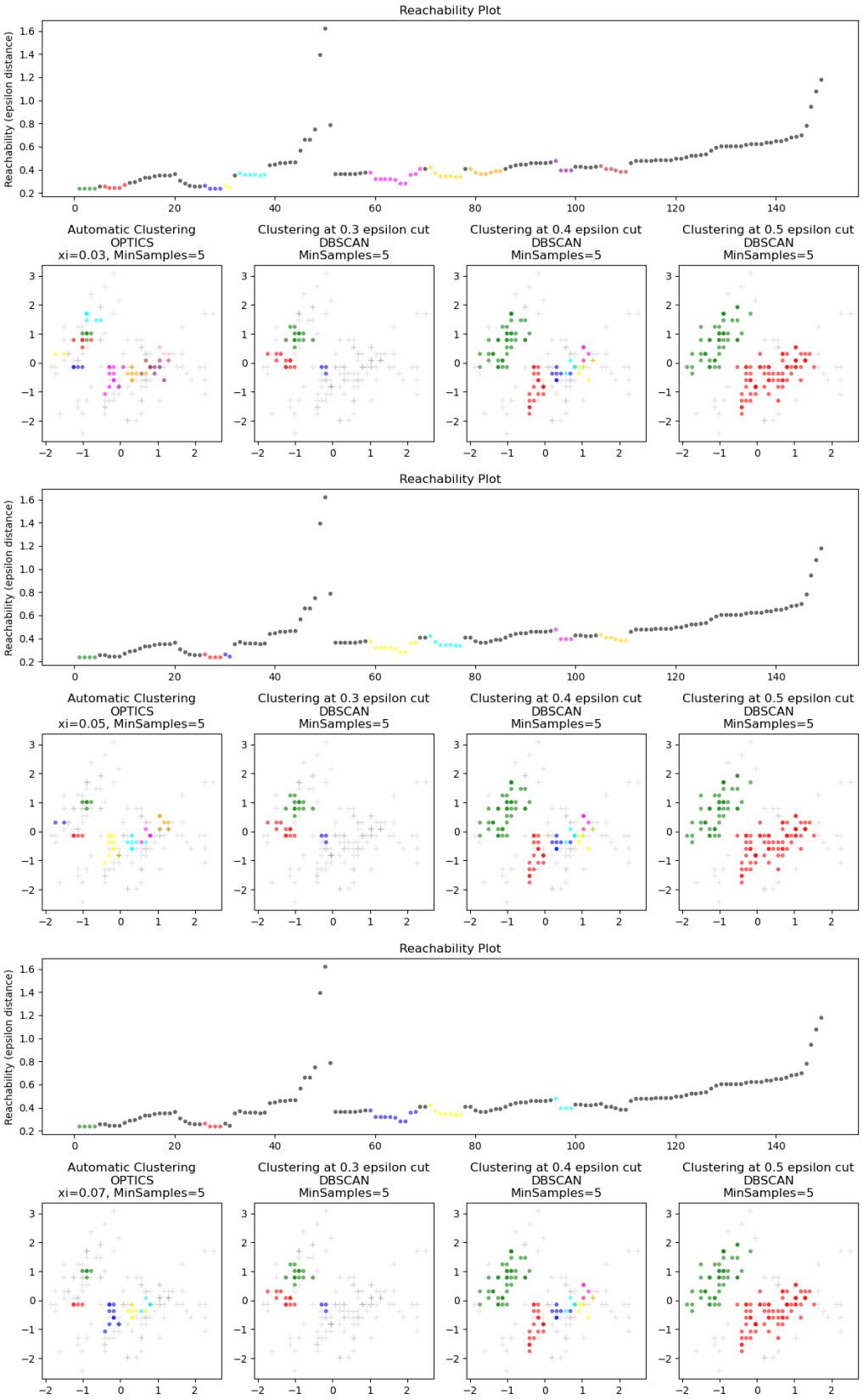



**Figure 2.** *Cont.*

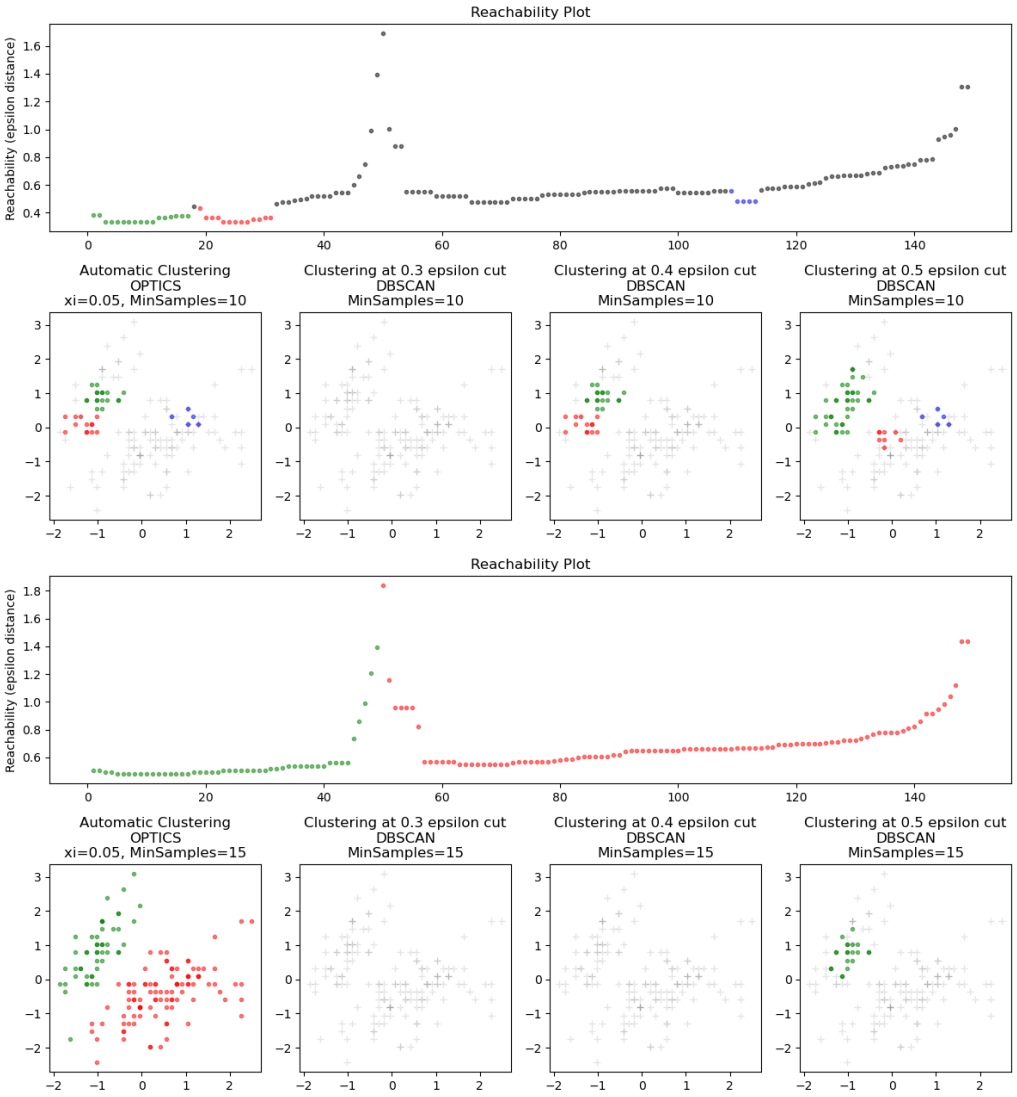



**Figure 2.** Different OPTICS clustering settings on Iris dataset (Example 1–5, observations within the same cluster are shown in the same colour).

It can be seen from the illustrations that under different settings, the results from the OPTICS clustering algorithm (regarding the variation of $\xi$) exhibit more effective clusters (in smaller groups and located in denser areas). This contrasts sharply with the outcomes of applying DBSCAN, where clusters are relatively larger across various densities (across Examples 1–3). DBSCAN may even generate few or no clusters when encountering a greater *MinSample* value (Example 4 with *MinSample* = 10 and Example 5 with *MinSample* = 15). The above empirical findings indicate that OPTICS offers stronger robustness when determining groups within a dense area of data points. More importantly, it facilitates clustering parameter optimisation in exploring given datasets. This differs from the existing GA-based approach [24] that utilises pre-defined hypercubes to depict the knowledge space. Applying a clustering algorithm avoids strict boundaries like hypercubes. Furthermore, the reachability feature in OPTICS introduces a new level of robustness on top of DBSCAN, which can determine tighter clusters, enabling the approach proposed herein to support the generalisation of the interpolated rules with more informative content.

### 2.3. Membership Frequency Weight Calculation and Interpolated Rule Promotion

Apart from the generic structure proposed for OPTICS-D-FRI, a key contribution of this work lies in providing a means to compute the membership frequency weights that are exploited to decide on whether an interpolated rule should be promoted to become a generalised rule, thereby enriching the existing sparse rule base. Algorithm 2 summarises the underlying procedure for such a computation.

To aid in understanding, a demonstration of how to compute frequency weights is provided here, based on a sparse rule base as indicated in Table 2. For simplicity, suppose that all fuzzy terms are represented by their respective representative values. This table shows that one of three membership functions (say, $M_0, M_1$ and $M_2$) may be taken by each of the three antecedent variables at a time and that there are three potential consequent values too. For classification tasks, the consequent values may be crisp, denoting three different classes (say, classes 0, 1, and 2). Thus, the first row of this table represents a fuzzy rule, stating that "If $x_1$ is $A_1$ and $x_2$ is $A_2$ and $x_3$ is $A_3$ then $y$ is 0", where fuzzy sets $A_1$, $A_2$, $A_3$, and $B_2$ are defined with certain membership functions $M_1$, $M_0$, and $M_1$ (whose representative values are 1, 0 and 1), respectively. The use of representative values herein helps simplify the explanation of the relevant calculations.

**Algorithm 2** Rule base frequency weight and rule scoring application.

**Start**:
**Input:** Sparse rule base $\mathcal{R}^S$,

$$R_i : \text{IF } x_1 \text{ is } A_{i,1} \text{ AND ... AND } x_N \text{ is } A_{i,N} \text{ THEN } y \text{ is } B_i,$$

where $i$ spans the set $\{1, \ldots, |\mathcal{R}^S|\}$ with $|\mathcal{R}^S|$ denoting the total count of rules present in the sparse rule base, $N$ indicates the overall number of these variables, and the notation $A_{i,j}$ indicates the linguistic value attributed to the variable $x_j$ within the rule $R_i$ (herein represented as a triangular fuzzy set $(a_{0,j}, a_{1,j}, a_{2,j})$).
**Output:** Frequency weights table and frequency score

**Step 1.** For each rule $R_i$ in $\mathcal{R}^S$:

Extract representative values $Rep(A_{i,j}) = (a_{0,j}, a_{1,j}, a_{2,j})/3$ from each antecedent and $Rep(B) \in \{0, 1, \cdots, C\}$ from the consequent (where $C$ is the total number of decision categories, e.g., the number of possible classes for classification problems).

**Step 2.** Identify all $q$ different rules per decision category $Rep(B) \in \{0, 1, \cdots, C\}$ and calculate the frequency by $1/n$ from each antecedent domain, where $n$ stands for the number of unique $Rep(A_{i,j})$ within $q$ different rules of each class, with $i \in \{1, \ldots, q\}$ in each domain $j \in \{1, \ldots, N\}$.

**Step 3.** Update the frequency weights table of each antecedent domain regarding each possible class $Rep(B) \in \{0, 1, \cdots, C\}$.

**Step 4.** Compute $Rep(A_{fj})$ on each interpolated rule and match against frequency weights $W_f = (w_1, w_2, \ldots, w_j)$ within the frequency weights table for each $Rep(B) \in \{0, 1, \cdots, C\}$.

**Step 5.** Sum of the frequency weights $S_{Rep(B)} = w_1 + w_2 + \cdots + w_j$ for each $Rep(B) \in \{0, 1, \cdots, C\}$, resulting in $Score_B = [S_{Rep(B)=0}, S_{Rep(B)=1}, \ldots, S_{Rep(B)=C}]$

**Step 6.** Take $Max(Score_B) = Max[S_{Rep(B)=0}, S_{Rep(B)=1}, \ldots, S_{Rep(B)=C}]$, and replace the interpolated consequent with the outcome of $Max(Score_B)$.
**End.**

For example, consider the two rules that each conclude with class 0 (i.e., the first and last one, where $Rep(B = 0)$). Both rules are associated with the second membership function ($Rep(A_1) = 1$). However, regarding the first antecedent variable, there is just one unique fuzzy set for it to take, so $1/1 = 1$ for this variable. Regarding the second antecedent variable, it involves two different fuzzy sets ($M_0$ and $M_1$), as depicted by $Rep(A_1)$ and $Rep(A_2)$, which gives $1/2 = 0.5$ for each unique variable. Generalising the intuition underlying this example, the frequency weights can be determined following the procedure in Algorithm 2.

The results are displayed in Table 3, again denoted using their representative values for fuzzy terms.

**Table 2.** Example sparse rule base represented with representative values.

| $Rep(A_1)$ | $Rep(A_2)$ | $Rep(A_3)$ | $Rep(B)$ |
|---|---|---|---|
| 1 | 0 | 1 | 0 |
| 2 | 1 | 1 | 2 |
| 0 | 2 | 0 | 1 |
| 2 | 0 | 1 | 2 |
| 1 | 1 | 1 | 0 |

From the above, again for illustration, suppose that after step (7), the interpolated rule taken from the pool for checking is $\mathcal{R}_i^S$: $[Rep(A_1) = 1, Rep(A_2) = 2, Rep(A_3) = 1, Rep(B) = 0]$ (whose interpretation in terms of a conventional product rule for classification is obvious). Matching it against the frequency weights computed from the existing sparse rule base leads to the following: $Score_B = [S_{Rep(B)=0} : 1+0+1 = 2, S_{Rep(B)=1} : 0+1+0 = 1, S_{Rep(B)=2} : 0+0+1 = 1]$. In this case, $Rep(B) = 0$ is associated with the highest score, while the selected rule is considered with the same outcome. Therefore, this selected interpolated rule is promoted, becoming a new rule in the sparse rule base.

**Table 3.** Frequency weights of sparse rule base.

| | $Rep(A_1)$ | | | $Rep(A_2)$ | | | $Rep(A_3)$ | | |
|---|---|---|---|---|---|---|---|---|---|
| $Rep(B)$ | $M_0$ | $M_1$ | $M_2$ | $M_0$ | $M_1$ | $M_2$ | $M_0$ | $M_1$ | $M_2$ |
| 0 | 0 | 1 | 0 | 0.5 | 0.5 | 0 | 0 | 1 | 0 |
| 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 |
| 2 | 0 | 0 | 1 | 0.5 | 0.5 | 0 | 0 | 1 | 0 |

## 3. Density-Based Dynamic Fuzzy Rule Interpolation

This section describes the proposed method, by employing OPTICS to construct a dynamic fuzzy rule interpolation, hereafter referred to as OPTICS-D-FRI for simplicity.

*Structure of Proposed Approach*

To provide an overview of the approach introduced herein, Figure 3 illustrates the structure of OPTICS-D-FRI. Throughout the subsequent procedures, steps are enumerated with a notation '($k$)', representing the sequence of operations involved.

Initiated with a given sparse rule base, denoted as $\mathcal{R}$, and an unseen observation (which has not been utilised to derive any outcome previously), the first step of the inference process tests whether any rule exists within $\mathcal{R}$ that can match the observation; if so, the respective rule is triggered to produce the outcome. In scenarios where the observation does not match any rule (2), the T-FRI (3) is employed to begin the rule interpolation process. The outcome, $\mathcal{R}_i^T$, is accumulated into a pool of interpolated rules. After accumulating a complete batch of observations in the interpolated pool (6), the OPTICS algorithm clusters the corresponding observation points (4). Then, the interpolated rules are categorised in alignment with the clustering outcomes of the observations (5–6). After that, the recurrence frequency of interpolated rules within each cluster is calculated. Those rules $\mathcal{R}_i^S : A_{m,1}, A_{m,2}, \ldots, A_{m,N}, B_z$ (with $m$ denoting the number of attributes in the rule antecedent space and $z$ representing the fuzzy consequent) of a frequency $Freq_{\mathcal{R}_i^S}$ surpassing the mode frequency $mode(Freq_{C_i})$ within the cluster $C_i$, are selected to progress onto the subsequent phase (7).

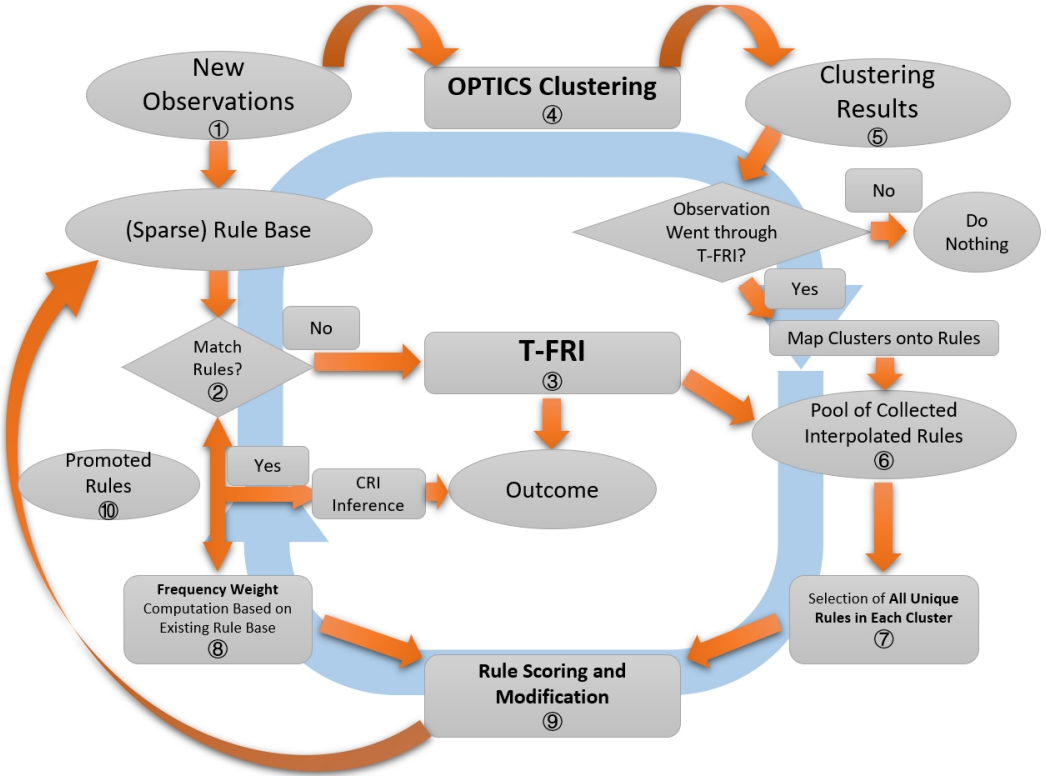



**Figure 3.** Proposed OPTICS-D-FRI framework.

Regarding every potential outcome $B_z$, stage (8) computes a linguistic attribute frequency weight based on the sparse rule base, represented by $(\mathcal{W}_{sparse}^{B_z} = [W_{A0}^0, W_{A1}^1, \ldots, W_{AN}^N])$, where $W_{AN}^N = [a_{N0}, a_{N1}, \ldots, a_{Nm}]$, and $a_{Nm}$ signifies the frequency weight of an attribute within the antecedent $A_N$ corresponding to $B_z$. The procedure following this calculates a score, $Score_B = [S_{B1}, S_{B2}, \ldots, S_{Bz}]$, for each consequent $B_z$ regarding the antecedent linguistic attributes of the rule $\mathcal{R}_i^S : A_{m,1}, A_{m,2}, \ldots, A_{m,N}$ (9). If the resultant score, $S_{Bz}$, for $B_z$ from the rule $\mathcal{R}_i^S : A_{m,1}, A_{m,2}, \ldots, A_{m,N}, B_z$ is the highest, this rule is selected (10) and promoted into the original sparse rule base, leading to an enriched rule base $\mathcal{P} = \mathcal{R} \cup \{\mathcal{R}_i^S\}$ (10). Otherwise, if the score $S_{Bz}$ does not stand out within $Score_B$, the consequent $B_z$ is substituted by $B_z = B_{MAX(Score_B)}$, and then the entire process iterates.

Note that the computational complexity of OPTICS and that of T-FRI [12] are $O_{OPTICS}(n^2)$ and $O_{T-FRI}(n^2)$, respectively. Therefore, the overall computational complexity of the proposed approach is $O_{OPTICS-D-FRI} = O_{OPTICS}(n^2) + O_{T-FRI}(n^2)$. This means that $O_{OPTICS-D-FRI} = O(n^2)$ (Algorithm 3).

**Algorithm 3** OPTICS-D-FRI Procedure.

**Require:** Sparse rule base $\mathcal{R}$, unseen observations
**Ensure:** Enriched or promoted rule base $\mathcal{P}$
Initialize rule base $\mathcal{R}$
Initialize interpolated rule pool $\mathcal{R}_i^T = \emptyset$
**for** each observation **do**
  **if** there exists a matching rule in $\mathcal{R}$ **then**
    Trigger and execute respective rule
  **else**
    Apply T-FRI to compute rule interpolation
    Add interpolated outcome to $\mathcal{R}_i^T$
  **end if**
**end for**
Cluster all unseen observations using the OPTICS algorithm, obtain clusters $C = \{C_1, C_2, \ldots, C_i\}$
**for** each cluster $C_i$ **do**
  Categorize interpolated rules with respect to $C_i$ of the related observation that generates each interpolated rule
  Select all unique rules in the cluster for the next phase
**end for**
**for** each potential outcome $B_z$ **do**
  Compute linguistic attribute frequency weight $\mathcal{W}_{sparse}^{B_z}$
  Calculate score $Score_B$ for each $B_z$
  **for** each rule in selected rules **do**
    **if** $S_{Bz}$ is highest in $Score_B$ **then**
      Promote the corresponding rule
      $\mathcal{P} = \mathcal{R} \cup \{\mathcal{R}_i^S\}$
    **else**
      Substitute $B_z$ with $B_{MAX(Score_B)}$
      $\mathcal{P} = \mathcal{R} \cup \{\mathcal{R}_i^S\}$
    **end if**
  **end for**
**end for**

## 4. Illustrative Example of OPTICS-D-FRI

To showcase the running process of the OPTICS-D-FRI algorithm, for simplicity, a trivial scenario (with a small dataset and a set of new observations) is assumed. However, this example illustrates the key steps of the algorithm, including data clustering with OPTICS, rule interpolation with T-FRI, and the rule promotion process, which takes into consideration different aspects, including rule uniqueness, rule base attribute frequency weights, and rule scoring.

**Dataset:**

| Observation | Feature 1 | Feature 2 | Class |
|---|---|---|---|
| 1 | 1.0 | 3.0 | A |
| 2 | 2.0 | 4.0 | A |
| 3 | 10.0 | 12.0 | B |
| 4 | 11.0 | 11.0 | B |
| 5 (New) | 2.0 | 3.5 | ? |
| 6 (New) | 1.5 | 3.2 | ? |
| 7 (New) | 10.5 | 12.5 | ? |
| 8 (New) | 9.5 | 11.5 | ? |
| 9 (New) | 1.8 | 2.9 | ? |
| 10 (New) | 2.2 | 4.1 | ? |
| 11 (New) | 10.2 | 11.8 | ? |
| 12 (New) | 9.8 | 11.6 | ? |
| 13 (New) | 2.1 | 3.3 | ? |
| 14 (New) | 1.7 | 3.6 | ? |

### *4.1. Dataset and Sparse Rule Base*

Consider a dataset with observations in a two-dimensional feature space and their corresponding classes. The sparse rule base (SRB) initially contains fuzzy rules covering certain sub-regions of the problem.

**Initial sparse rule base (SRB):**

- Rule 1: IF Feature 1 is "low" AND Feature 2 is "medium" THEN Class A.
- Rule 2: IF Feature 1 is "high" AND Feature 2 is "high" THEN Class B.

### *4.2. Clustering with OPTICS and T-FRI Application*

Applying the OPTICS algorithm for the clustering of data objects and T-FRI over new observations (which are not matched by the rules given) leads to:

**Clusters identified by OPTICS:**

- Cluster 1: Observations 1, 2, 5, 6, 9, 10, 13, 14 (Close to Class A)
- Cluster 2: Observations 3, 4, 7, 8, 11, 12 (Close to Class B)

**T-FRI interpolated rules:**

- Interpolated rules due to observations 5, 6, 9, 10, 13, 14 go to Class A.
- Interpolated rules due to observations 7, 8, 11, 12 go to Class B.

### *4.3. Rule Promotion Considering Uniqueness and Attribute Frequency Weight*

The rule promotion process considers the uniqueness of the patterns identified in the clusters and incorporates attribute frequency weights and scoring.

**Promotion criteria:**

- Promote rules that are unique within a cluster.
- Use attribute frequency weights from the rule base to evaluate the score of each individual interpolated rule.

**Promoted rules:**

- IF Feature 1 is "medium" AND Feature 2 is "high" THEN Class A.
- IF Feature 1 is "high " AND Feature 2 is "very high" THEN Class B.

**Enriched rule base:**

- Rule 1: IF Feature 1 is "low" AND Feature 2 is "medium" THEN Class A.
- Rule 2: IF Feature 1 is "high" AND Feature 2 is "high" THEN Class B.
- Rule 3: IF Feature 1 is "medium" AND Feature 2 is "high" THEN Class A (New).
- Rule 4: IF Feature 1 is "high " AND Feature 2 is "very high" THEN Class B (New).

## 5. Experimental Investigations

This section explains the experimental setup and discusses the results of applying the proposed OPTICS-D-FRI.

### *5.1. Datasets and Experimental Setup*

As an initial attempt toward the development of a comprehensive dynamic fuzzy rule interpolative reasoning tool, five classical public datasets, namely Iris (iris flowers data between three species), Yeast (genetic, proteomic, and cellular data used for biological research), Banana (Kaggle example), Glass (analysis of different types of glass for forensic purposes), and Tae (evaluation of teaching assistants' performance based on various attributes), are herein employed to validate the performance of the methodology presented. Note that given the restriction on the size of the Iris dataset, which contains just 150 observations, it is necessary to incorporate data simulation techniques to ensure a more robust assessment of the proposed approach. For this purpose, the Iris dataset has been artificially expanded using the 'faux' package [30] in the R environment. Detailed descriptions of all five datasets are presented in Table 4.

**Table 4.** Dataset descriptions.

| | Iris (Expanded) | Yeast | Banana | Glass | Tae |
|---|---|---|---|---|---|
| Features | 4 | 9 | 2 | 9 | 6 |
| Classes | 3 | 10 | 2 | 7 | 3 |
| Sample Size | 1500 | 1484 | 5300 | 214 | 151 |

For each application problem (i.e., dataset), the dataset is partitioned into three distinct categories. This partition is randomly carried out, with the exception that the sizes of the partitioned sub-datasets are pre-specified. One that comprises 50% of the total data is used to form a dense rule base for each problem, denoted as $\mathcal{D}$. The resulting rule base covers the majority of the problem domain and is intended to be the gold standard for subsequent comparative studies when a sparse rule base is used. Adopting a conventional data-fitting algorithm as delineated in [31] facilitates the rule generation process. More advanced rule induction techniques (e.g., the method of [32,33]) may be utilised as alternatives for enhanced outcomes if preferred.

Note that in the experimental setup, datasets are divided into three parts: The first part (50%), namely the training set, is used to generate a dense rule base with a data-fitting algorithm. Subsequently, after the formation of a dense rule base $\mathcal{D}$, a strategic rule reduction is performed by randomly dropping 30% of the rules from the dense rule base to create a sparse rule base, with the remaining 70% being regarded as the sparse rule base $\mathcal{R}$. The second part of the data (25%) is then utilised to generate a pool of interpolated rules by running T-FRI, which occurs when an observation does not match any rule in the sparse rule base. After the pool of interpolated rules is obtained, the D-FRI approach is initiated to select rules from this emerging pool to create a 'promoted' rule base $\mathcal{P}$. The final part of the data (25%) is used to evaluate the performance of both conventional T-FRI and D-FRI. While T-FRI always relies on the sparse rule base, D-FRI exploits all previously promoted rules as well. If the enriched rule base still fails to derive a result, given an unmatched observation, D-FRI then defaults to running T-FRI as usual to derive the consequent, and this newly interpolated rule is saved into a new pool for a potential new iteration. However, as an initial formalised proposal of density-based D-FRI, only one iteration is presently performed. This setup allows for a direct comparison of accuracy between the D-FRI-promoted rules and the conventional T-FRI rules. Additionally, the dynamically promoted rule base with weights, $\mathcal{P}_{\mathcal{D}}$, denoted by OPTICS-D-FRI(DW), is also evaluated using the third sub-dataset.

A $10 \times 10$ cross-validation is implemented on the test datasets to evaluate the proposed OPTICS-D-FRI algorithm. For easy comparison, metrics such as average accuracy and associated standard deviation regarding the employment of conventional rule-firing with just the sparse rule base (SRB), T-FRI, OPTICS-D-FRI, and the dynamically weighted OPTICS-D-FRI are presented. Within these tables, the better performances are shown in bold font.

### *5.2. Experimental Results*

The results of performance evaluation are presented in the following five tables (Tables 5–9), with each corresponding to a given dataset. According to these results, the proposed OPTICS-D-FRI methodology proficiently clusters rules for unmatched observations. The introduction of the frequency weight method enables a dynamic equilibrium between T-FRI interpolated rules and those from the original sparse rule base, demonstrating a capability to simulate human reasoning. The inference outcomes from using the present approach improve significantly over those attainable by running conventional FRI methods given just a sparse rule base. The performance is strengthened by utilising the frequency weights derived from the sparse rule base. The promoted rules exhibit an accuracy that is at least equal to, and often higher than, the interpolated rules.

**Table 5.** The 10 × 10 cross-validation on the Expanded Iris dataset, with four membership functions defined for each antecedent variable.

| | *OPTICS: xi = 0.08 MinSample = 5* | *OPTICS: xi = 0.05 MinSample = 5* | *OPTICS: xi = 0.05 MinSample = 10* |
|---|---|---|---|
| | **Expanded Iris (Accuracy ± Standard Deviation)** | | |
| **SRB** | 0.474 ± 0.147 | 0.52 ± 0.084 | 0.556 ± 0.125 |
| **T-FRI** | 0.809 ± 0.099 | 0.788 ± 0.108 | 0.747 ± 0.142 |
| **OPTICS-D-FRI** | **0.825 ± 0.072** | **0.813 ± 0.082** | **0.789 ± 0.097** |
| **OPTICS-D-FRI(DW)** | **0.818 ± 0.072** | **0.811 ± 0.069** | **0.787 ± 0.096** |
| | **Rule Base Coverage (Avg.)** | | |
| **SRB** | 0.54 | 0.58 | 0.63 |
| **OPTICS-D-FRI** | **0.88** | **0.93** | **0.86** |

**Table 6.** The 10 × 10 cross-validation on the Yeast dataset, with four membership functions defined for each antecedent variable.

| | *OPTICS: xi = 0.08 MinSample = 5* | *OPTICS: xi = 0.05 MinSample = 5* | *OPTICS: xi = 0.08 MinSample = 10* |
|---|---|---|---|
| | **Yeast (Accuracy ± Standard Deviation)** | | |
| **SRB** | 0.238 ± 0.039 | 0.249 ± 0.045 | 0.23 ± 0.054 |
| **T-FRI** | 0.281 ± 0.073 | 0.299 ± 0.061 | 0.308 ± 0.034 |
| **OPTICS-D-FRI** | **0.298 ± 0.064** | **0.314 ± 0.052** | **0.33 ± 0.043** |
| **OPTICS-D-FRI(DW)** | **0.295 ± 0.062** | **0.312 ± 0.049** | **0.325 ± 0.04** |
| | **Rule Base Coverage (Avg.)** | | |
| **SRB** | 0.56 | 0.57 | 0.52 |
| **OPTICS-D-FRI** | **0.82** | **0.81** | **0.75** |

**Table 7.** The 10 × 10 cross-validation on the Banana dataset, with four membership functions defined for each antecedent variable.

| | *OPTICS: xi = 0.05 MinSample = 10* | *OPTICS: xi = 0.05 MinSample = 20* | *OPTICS: xi = 0.1 MinSample = 10* |
|---|---|---|---|
| | **Banana (Accuracy ± Standard Deviation)** | | |
| **SRB** | 0.354 ± 0.096 | 0.36 ± 0.123 | 0.335 ± 0.113 |
| **T-FRI** | 0.466 ± 0.099 | 0.545 ± 0.083 | 0.454 ± 0.07 |
| **OPTICS-D-FRI** | **0.515 ± 0.062** | **0.565 ± 0.054** | **0.506 ± 0.047** |
| **OPTICS-D-FRI(DW)** | **0.511 ± 0.0614** | **0.564 ± 0.053** | **0.511 ± 0.047** |
| | **Rule Base Coverage (Avg.)** | | |
| **SRB** | 0.57 | 0.56 | 0.54 |
| **OPTICS-D-FRI** | **0.99** | **0.98** | **0.98** |

**Table 8.** The 10 × 10 cross-validation on the Glass dataset, with four membership functions defined for each antecedent variable.

| | *OPTICS: xi = 0.3 MinSample = 2* | *OPTICS: xi = 0.3 MinSample = 3* | *OPTICS: xi = 0.1 MinSample = 2* |
|---|---|---|---|
| | **Glass (Accuracy ± Standard Deviation)** | | |
| **SRB** | 0.207 ± 0.064 | 0.209 ± 0.05 | 0.179 ± 0.063 |
| **T-FRI** | 0.369 ± 0.068 | 0.323 ± 0.06 | 0.41 ± 0.04 |
| **OPTICS-D-FRI** | **0.385 ± 0.066** | **0.353 ± 0.057** | **0.414 ± 0.039** |
| **OPTICS-D-FRI(DW)** | **0.382 ± 0.065** | **0.349 ± 0.053** | **0.411 ± 0.038** |
| | **Rule Base Coverage (Avg.)** | | |
| **SRB** | 0.36 | 0.38 | 0.32 |
| **OPTICS-D-FRI** | **0.53** | **0.46** | **0.57** |

**Table 9.** The 10 × 10 cross-validation on the Tae dataset, with four membership functions defined for each antecedent variable.

| | ***OPTICS: xi = 0.05 MinSample = 2*** | ***OPTICS: xi = 0.05 MinSample = 4*** | ***OPTICS: xi = 0.02 MinSample = 2*** |
|---|---|---|---|
| | **Tae (Accuracy ± Standard Deviation)** | | |
| **SRB** | 0.2 ± 0.046 | 0.202 ± 0.024 | 0.189 ± 0.051 |
| **T-FRI** | 0.328 ± 0.053 | 0.341 ± 0.087 | 0.327 ± 0.055 |
| **OPTICS-D-FRI** | **0.336 ± 0.045** | **0.361 ± 0.082** | **0.343 ± 0.053** |
| **OPTICS-D-FRI(DW)** | **0.338 ± 0.045** | **0.363 ± 0.08** | **0.347 ± 0.049** |
| | **Rule Base Coverage (Avg.)** | | |
| **SRB** | 0.33 | 0.38 | 0.37 |
| **OPTICS-D-FRI** | **0.48** | **0.51** | **0.52** |

### *5.3. Discussion*

In implementing both OPTICS and T-FRI, the Euclidean distance metric is uniformly employed for distance measurement. Within the OPTICS algorithm, the parameter denoted as *MinSample* dictates the general size of the resulting clusters, essentially controlling the number of samples assigned to a specific cluster. Prior to parameter configuration, conducting an exploratory analysis of the problem domain (or dataset) is required. As an illustration, the Iris dataset has four features with distinctive clustering boundaries spanning four distinct classes. Conversely, the Yeast dataset is distributed across nine features, pairing with ten classes. Exploring the dataset and investigating the clustering performance is, therefore, necessary instead of randomly setting up OPTICS parameters; domain-specific insights can facilitate optimal cluster formation.

In empirical studies, it is observed that the *MinSample* parameter controls the size of the clusters. Consequently, in the present preliminary experimental investigations, the other parameter, *MinPts* (the minimum number of samples to form a cluster), is fixed at two. Furthermore, the number of membership functions is equally assigned onto the span of the antecedent space without any designated design, which may lead to the overall low accuracy of these datasets. However, this does not adversely affect the present study since the relative performance is considered among different approaches. An optimised domain partition can be expected to produce better outcomes. From the results obtained, the promoted rules not only enlarge the coverage of the original sparse rule base but also enable the inference system to attain better accuracy compared to traditional T-FRI.

The dynamic fuzzy rule interpolation process enhances the system's applicability by allowing the rule base to adapt and evolve in response to new and unmatched observations. This adaptability is crucial in dealing with problems within a dynamic and complex environment, where the underlying data patterns can change over time. Such enrichment of the (sparse) rule base will enable future pattern-matching-based rule-firing processes to be more efficiently implemented, without the need to go through computationally more expensive interpolated processes, especially for those similar (and currently unmatched) observations.

Upon closer examination of the experimental results, it can be observed that certain promoted rules align with some of those rules (intentionally removed for testing purposes) from the original dense rule base. This observation underscores the effectiveness of the proposed approach. This suggests that promoted rules indeed help improve the efficiency of the inference system, particularly when observations close to previously unmatched data are introduced. It allows the execution of reasoning under such conditions with simple pattern-matching, subsequently reducing the need for FRI processes and, hence, saving computational efforts while strengthening reasoning accuracy. Furthermore, merging a dynamic attribute weighting mechanism with the traditional T-FRI method, which accounts for the newly promoted rules in frequency weight calculation, offers considerable promise in enhancing the performance of a fuzzy rule-based system.

## 6. Conclusions

This paper proposes an initial approach for creating a dynamic FRI framework by integrating the OPTICS clustering algorithm with the conventional T-FRI. It works by extracting information embedded within interpolated rules in relation to new and unmatched observations. The paper also introduces a novel method for determining populated regions by clustering the new observations and annotating the group of interpolated rules, thereby eliminating the need for predefined hypercubes. The density-based method evaluates the significance of each interpolated rule by considering how frequently similar situations have occurred. By promoting interpolated rules resulting from more frequently incurred data points, the system can focus on more relevant and generalisable knowledge, thereby improving the quality and reliability of future inference outcomes. For any future similar (and currently unmatched) observations, the dynamic rule promotion process will discount those otherwise required interpolation computations, thus significantly increasing the efficiency of the overall rule-based inference process. In addition, the intuitive approach taken for frequency weight scoring and rule modification mechanisms provides an effective way to evaluate and reason with fuzzy rules. Notably, the proposed OPTICS-D-FRI demonstrates enhanced performance over the traditional FRI in terms of reasoning accuracy.

Being an initial attempt to combine OPTICS and T-FRI for the implementation of a system within the proposed dynamic FRI framework, this study also identifies several limitations for further refinement. For instance, the current method requires specifically predefined parameters, including $MinSample$ and $\xi$ for OPTICS. Investigating a mechanism to automatically evaluate observations and determine such parameters remains active research. As an initial realisation of a formalised D-FRI mechanism, the stability of its performance for online applications forms an interesting issue for further investigation. Reinforcing the frequency weight scoring based on the fuzzy values is also an important step to take further. In addition, the foundational framework combining OPTICS is not strictly bound to the T-FRI method used in this study. Thus, a possible future direction is to examine the integration of alternative FRI methods and to evaluate their performance.

One motivation for developing the present work is to obtain a computational technique useful for solving real-world medical problems (e.g., for breast cancer detection based on mammography [34]). This requires handling high-dimensional data, for which recent advancements in FRI literature, especially those utilising attribute-weighted rules [28] and exploring rule base structures, may be helpful if reinforced. However, type-2 fuzzy systems [35], with the ability to manage membership functions that are fuzzy, offer a more complex framework for modelling the ambiguities and uncertainties. The advancement in type-2 fuzzy systems [36] sets a helpful foundation for seeking solutions to such high-dimensional problems. This may draw inspiration from the existing research on higher-order FRI [37]. For image-based analysis, in particular, the recent developments in fuzzy neural networks have become potential alternatives [38].

Another exciting future exploration involves the assessment of dynamically weighted attribute frequency estimations, particularly with the inclusion of newly promoted rules. Broader empirical investigations involving more complex data (again, for mammography risk analysis as an example) will also remove the present limitations in conducting empirical studies that only involve a small number of classical datasets. Furthermore, if many rules are promoted over the dynamic learning process, this work may encounter the likelihood that the rule base becomes unnecessarily dense over time. Thus, developing a closed loop for the D-FRI model, allowing the dynamic FRI system to maintain its efficiency without losing its self-updating ability, requires further research.

**Author Contributions:** Conceptualisation, J.L. and Q.S.; methodology, J.L. and C.S.; software, J.L.; validation, C.S., J.L., and Q.S.; formal analysis, J.L.; investigation, J.L. and C.S.; resources, J.L. and C.S.; writing—original draft preparation, J.L.; writing—review and editing, J.L., C.S., and Q.S.; visualization, J.L.; supervision, C.S. and Q.S.; project administration, C.S.; funding acquisition, C.S. and Q.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** The data presented in this study are public datasets.

**Acknowledgments:** The authors are grateful to the anonymous reviewers for their helpful comments on the original conference paper, which was recommended for expanded consideration in this special issue. Thanks also go out to the anonymous reviewers who provided constructive feedback on the previous versions of the expanded submission, which has helped improve the quality of this work significantly. The first author is supported in part by an Aberystwyth University PhD scholarship.

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

## Abbreviations

The following abbreviations are used in this manuscript:

| | |
|---|---|
| FRI | fuzzy rule interpolation |
| D-FRI | dynamic fuzzy rule interpolation |
| T-FRI | transformation-based fuzzy rule interpolation |
| OPTICS | ordering points to identify the clustering structure |
| GA | genetic algorithm |
| PSO | particle swarm optimisation |
| DBSCAN | density-based spatial clustering of applications with noise |
| OPTICS-D-FRI | OPTICS assisted dynamic fuzzy rule interpolation |

Several notations are used to describe the concepts and methodologies. For clarity and ease of understanding, the key notations used throughout the paper are listed below:

| | |
|---|---|
| $\mathcal{R}$ | represents the sparse rule base |
| $\mathcal{D}$ | denotes the dense rule base |
| $R_i$ | indicates a specific rule within the rule base |
| $A_{i,j}$ | symbolises the linguistic value of the $j$th antecedent in the $i$th rule |
| $B_i$ | represents the consequent of the $i$th rule |
| $O$ | refers to an observation |
| $\mathcal{R}_i^T$ | denotes an interpolated rule |
| $\mathcal{P}$ | indicates the promoted rule base |
| $C_i$ | represents a cluster identified by OPTICS |
| $\xi$ | parameter in OPTICS defining the steepness for cluster boundary |
| $MinSample$ | minimum number of samples in a cluster for OPTICS |
| $\mathcal{P}_\mathcal{D}$ | denotes the dynamically promoted rule base with frequency weights |
| $Score_B$ | the scoring mechanism for evaluating rules |
| $\omega_{i,j}$ | normalised displacement factor in T-FRI |

## References

1. Kóczy, L.; Hirota, K. Ordering, distance and closeness of fuzzy sets. *Fuzzy Sets Syst.* **1993**, *59*, 281–293. [CrossRef]
2. Koczy, L.T.; Hirota, K. Size reduction by interpolation in fuzzy rule bases. *IEEE Trans. Syst. Man Cybern. Part B (Cybern.)* **1997**, *27*, 14–25. [CrossRef] [PubMed]
3. Dubois, D.; Prade, H. On fuzzy interpolation. *Int. J. Gen. Syst.* **1999**, *28*, 103–114. [CrossRef]
4. Kóczy, L.; Hirota, K. Approximate reasoning by linear rule interpolation and general approximation. *Int. J. Approx. Reason.* **1993**, *9*, 197–225. [CrossRef]
5. Kóczy, L.; Hirota, K. Interpolative reasoning with insufficient evidence in sparse fuzzy rule bases. *Inf. Sci.* **1993**, *71*, 169–201. [CrossRef]
6. Kóczy, L.; Hirota, K. Rule interpolation by $\alpha$-level sets in fuzzy approximate reasoning. *J. BUSEFAL Automne URA-CNRS* **1991**, *46*, 115–123.

7. Gedeon, T.; Kóczy, L. Conservation of fuzziness in rule interpolation. *Symp. New Trends Control. Large Scale Syst.* **1996**, *1*, 13–19.
8. Hsiao, W.H.; Chen, S.M.; Lee, C.H. A new interpolative reasoning method in sparse rule-based systems. *Fuzzy Sets Syst.* **1998**, *93*, 17–22. [CrossRef]
9. Kóczy, L.T.; Hirota, K.; Gedeon, T.D. Fuzzy Rule Interpolation by the Conservation of Relative Fuzziness. *J. Adv. Comput. Intell. Intell. Inform.* **2000**, *4*, 95–101. [CrossRef]
10. Baranyi, P.; Gedeon, T.D.; Koczy, L. Rule interpolation by spatial geometric representation. *Proc. IPMU* **1996**, *96*, 483–488.
11. Bouchon-Meunier, B.; Marsala, C.; Rifqi, M. Interpolative reasoning based on graduality. In Proceedings of the Ninth IEEE International Conference on Fuzzy Systems, FUZZ-IEEE 2000 (Cat. No. 00CH37063), IEEE, San Antonio, TX, USA, 7–10 May 2000; Volume 1, pp. 483–487.
12. Huang, Z.; Shen, Q. Fuzzy interpolation and extrapolation: A practical approach. *IEEE Trans. Fuzzy Syst.* **2008**, *16*, 13–28. [CrossRef]
13. Jenei, S.; Klement, E.P.; Konzel, R. Interpolation and extrapolation of fuzzy quantities–The multiple-dimensional case. *Soft Comput.* **2002**, *6*, 258–270. [CrossRef]
14. Johanyák, Z.C.; Kovács, S. Fuzzy rule interpolation by the least squares method. In Proceedings of the 7th International Symposium of Hungarian Researchers on Computational Intelligence (HUCI 2006), Budapest, Hungary, 24–25 November 2006; Volume 2425, pp. 495–506.
15. Lee, L.W.; Chen, S.M. Fuzzy interpolative reasoning for sparse fuzzy rule-based systems based on the ranking values of fuzzy sets. *Expert Syst. Appl.* **2008**, *35*, 850–864. [CrossRef]
16. Chen, S.M.; Chang, Y.C. Weighted fuzzy rule interpolation based on GA-based weight-learning techniques. *IEEE Trans. Fuzzy Syst.* **2011**, *19*, 729–744. [CrossRef]
17. Das, S.; Chakraborty, D.; Kóczy, L. Forward and backward fuzzy rule base interpolation using fuzzy geometry. *Iran. J. Fuzzy Syst.* **2023**, *20*, 127–146.
18. Åström, K.J.; Wittenmark, B. *Adaptive Control*; Courier Corporation: North Chelmsford, MA, USA, 2013.
19. Liu, B.D.; Chen, C.Y.; Tsao, J.Y. Design of adaptive fuzzy logic controller based on linguistic-hedge concepts and genetic algorithms. *IEEE Trans. Syst. Man Cybern. Part B (Cybern.)* **2001**, *31*, 32–53.
20. Mohan, S.; Bhanot, S. Comparative study of some adaptive fuzzy algorithms for manipulator control. *Int. J. Comput. Intell.* **2006**, *3*, 303–311.
21. Wang, Y.; Deng, H.; Chen, Z. Adaptive fuzzy logic controller with rule-based changeable universe of discourse for a nonlinear MIMO system. In Proceedings of the 5th International Conference on Intelligent Systems Design and Applications (ISDA'05), IEEE, Warsaw, Poland, 8–10 September 2005; pp. 8–13.
22. Labiod, S.; Boucherit, M.S.; Guerra, T.M. Adaptive fuzzy control of a class of MIMO nonlinear systems. *Fuzzy Sets Syst.* **2005**, *151*, 59–77. [CrossRef]
23. Nayak, D.; Ray, K.; Kar, T.; Mohanty, S.N. Fuzzy Rule Based Adaptive Block Compressive Sensing for WSN Application. *Mathematics* **2023**, *11*, 1660. [CrossRef]
24. Naik, N.; Diao, R.; Shen, Q. Dynamic fuzzy rule interpolation and its application to intrusion detection. *IEEE Trans. Fuzzy Syst.* **2017**, *26*, 1878–1892. [CrossRef]
25. Ester, M.; Kriegel, H.P.; Sander, J.; Xu, X. A density-based algorithm for discovering clusters in large spatial databases with noise. *KDD* **1996**, *96*, 226–231.
26. Sander, J.; Ester, M.; Kriegel, H.P.; Xu, X. Density-based clustering in spatial databases: The algorithm gdbscan and its applications. *Data Min. Knowl. Discov.* **1998**, *2*, 169–194. [CrossRef]
27. Ankerst, M.; Breunig, M.M.; Kriegel, H.P.; Sander, J. OPTICS: Ordering points to identify the clustering structure. *ACM Sigmod Rec.* **1999**, *28*, 49–60. [CrossRef]
28. Li, F.; Shan, C.; Li, Y.; Shen, Q. Interpretable mammographic mass classification with fuzzy interpolative reasoning. *Knowl. Based Syst.* **2020**, *191*, 105279. [CrossRef]
29. Fisher, R.A. The use of multiple measurements in taxonomic problems. *Ann. Eugen.* **1936**, *7*, 179–188. [CrossRef]
30. DeBruine, L. *Faux: Simulation for Factorial Designs*; R Package Version 1.1.0; R Core Team: Vienna, Austria, 2021. [CrossRef]
31. Wang, L.X.; Mendel, J. Generating fuzzy rules by learning from examples. *IEEE Trans. Syst. Man Cybern.* **1992**, *22*, 1414–1427. [CrossRef]
32. Niu, J.; Chen, D.; Li, J.; Wang, H. Fuzzy rule-based classification method for incremental rule learning. *IEEE Trans. Fuzzy Syst.* **2021**, *30*, 3748–3761. [CrossRef]
33. Sanz, J.A.; Bustince, H. A wrapper methodology to learn interval-valued fuzzy rule-based classification systems. *Appl. Soft Comput.* **2021**, *104*, 107249. [CrossRef]
34. Dounis, A.; Avramopoulos, A.N.; Kallergi, M. Advanced Fuzzy Sets and Genetic Algorithm Optimizer for Mammographic Image Enhancement. *Electronics* **2023**, *12*, 3269. [CrossRef]
35. Mendel, J.M. Type-2 fuzzy sets and systems: an overview. *IEEE Comput. Intell. Mag.* **2007**, *2*, 20–29. [CrossRef]
36. Mohammadzadeh, A.; Zhang, C.; Alattas, K.A.; El-Sousy, F.F.; Vu, M.T. Fourier-based type-2 fuzzy neural network: Simple and effective for high dimensional problems. *Neurocomputing* **2023**, *547*, 126316. [CrossRef]

37. Chen, C.; MacParthaláin, N.; Li, Y.; Price, C.; Quek, C.; Shen, Q. Rough-fuzzy rule interpolation. *Inf. Sci.* **2016**, *351*, 1–17. [CrossRef]
38. Kazemi, M.A.; Pa, M.; Uddin, M.N.; Rezakazemi, M. Adaptive neuro-fuzzy inference system based data interpolation for particle image velocimetry in fluid flow applications. *Eng. Appl. Artif. Intell.* **2023**, *119*, 105723. [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.