Article

Not peer-reviewed version

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

Jinle Lin , Changjing Shang * , Qiang Shen

Posted Date: 19 October 2023

doi: 10.20944/preprints202310.1248.v1

Keywords: Dynamic Rule Interpolation; Fuzzy Rule Interpolation; Density-Based Rule Promotion doi:10.20944/preprints202310.1248.v1

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: Fuzzy Rule Interpolation (FRI) offers an innovative methodology for inferring outcomes from data points that match no established rules in a sparse fuzzy rule base. Generally, a traditional fuzzy rule-based system suffers from the inability to perform inference while facing an unmatched observation, given a sparse fuzzy rule base that is derived from insufficient data or inadequate human experience and that fails to cover parts of a specific problem domain. Fortunately, starting with a sparse rule base, Fuzzy Rule Interpolation (FRI) can help to formulate interpolated rules specifically in instances where observations fail to activate any existing rule. An interpolated rule gives a result called consequent, which is related to the unmatched observation so that the inference system will not fail to generate an outcome corresponding to the observation as the output. Conventionally, these interpolated rules will be discarded after obtaining outcomes. However, valuable information may be potentially embedded within those discarded rules addressing the limitation regarding the coverage of the original sparse knowledge space. Moreover, without exploiting the potential value of the interpolated rules, the fuzzy inference system based on a sparse rule base will never improve its efficiency and robustness in the long run. Therefore, this underscores the necessity for Dynamic Fuzzy Rule Interpolation (D-FRI), which aims to collate interpolated rules that cover particular knowledge spaces absent in the sparse rule base. Intuitively, those frequently appeared interpolated rules are considered valuable because future observations are more likely to hit the respective problem areas. Promoting selected interpolated rules to the sparse rule base can enhance the overall coverage of the knowledge space and increase the inference efficiency over time. This paper introduces a framework for Dynamic Fuzzy Rule Interpolation that integrates the widely used Transformation-based Fuzzy Rule Interpolation (T-FRI) with an effective form of clustering of interpolated results. Experimental findings validate the effectiveness of this approach.

Keywords: Dynamic Rule Interpolation; Fuzzy Rule Interpolation; Density-Based Rule Promotion

---

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. Here, a sparse rule base means that 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 [1–5], when the rule base is incomplete and especially when it can dynamically update itself.

In general, there exist two categories of Fuzzy Rule Interpolation (FRI) methodologies. The first approach is centred on the utilisation of α-cuts, whereby the α-cut of an interpolated consequent is calculated from the α-cuts of its antecedents, as outlined by the Resolution Principle [6–8]. The second methodology employs the concept of similarity of fuzzy value and analogical reasoning [9–16]. Specifically, within this latter category, an FRI algorithm generates intermediate fuzzy rules based

 doi:10.20944/preprints202310.1248.v1

on the principle of similarity. Among such methods, Transformation-Based Fuzzy Rule Interpolation (T-FRI) is the most popular [11,12]. 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 [17–21]. 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. Approaches designed for sparse rule-based dynamic fuzzy rule interpolation are unfortunately rather limited. One such approach is fundamentally framed within a predefined quantity space of domain variables, called hyper-cubes [22]. This technique employs K-means clustering and genetic algorithms (GA) to sort through hyper-cubes of interpolated rules. Subsequently, it generates a new rule from the selected hyper-cube 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 (GA) 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) [23] and Ordering Points To Identify the Clustering Structure (OPTICS) [24] exist. Based on recent findings from applying Harmony Search as an optimisation method and DBSCAN as a clustering method [25,26], the density-based rule promotion approach is proven effective. 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 has the capability to filter out noise data, effectively excluding noise from the constructed clusters. When employed to select rules from a pool of interpolated rules, OPTICS can help to 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 more reliable for referencing in a dynamic system. Inspired by this observation, the present work applies OPTICS to implement a novel approach to D-FRI.

This paper is structured as follows. Section 2 introduces the core computational mechanism of T-FRI and that of OPTICS. Section 3 illustrates the proposed OPTICS-assisted density-based dynamic fuzzy rule interpolation framework. Section 4 provides experimental results over different data sets. Section 4.3 presents a discussion regarding the experimental results. Finally, Section 5 concludes the paper and outlines the potential improvements that would be interesting to be made in 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 for 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 [11,12]. 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 $\{1, \dots, |\mathcal{R}|\}$, with the term $|\mathcal{R}|$ signifying the total count of rules present in the rule base. In the given domain, each x_j with $j \in \{1, \dots, N\}$ 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 $A_{O,1}, \dots, A_{O,j}$ with $j \in \{1, \dots, N\}$. 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 an observation fails to match any rules within the sparse rule base.

Closest Rules Determination. 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 $A_j^{\dagger}$ in order 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 δ_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 ω_{B_i} and shift δ_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 below:

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

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)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 October 2023 doi:10.20944/preprints202310.1248.v1

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 [24], builds upon the foundation of DBSCAN [23], 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, which is measured by the parameter ξ, 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 ε is fixed; in OPTICS, ε 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.

Table 1. Concepts and parameters of OPTICS clustering algorithm.

Concept	Definition
ξ	Parameter used to specify the steepness that defines a cluster in reachability plot, with clusters determined as regions in reachability plot that exhibit a steep decrease, followed by a gentler increase in reachability distance.
MinPts	Minimum number of points needed to form a dense region, signifying that at least MinPts points are contained within its ε-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 MinPts points in its neighbourhood.
Ordering	Sequence in which points are processed, resulting in cluster ordering rather than explicit cluster assignments.

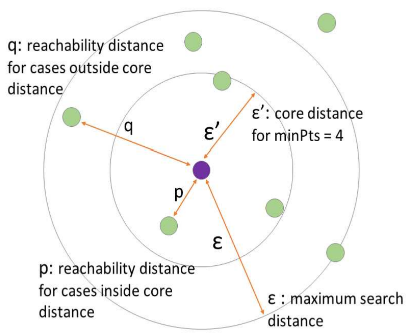

Figure 1. Illustration for OPTICS.

More formally, Algorithm 1 outlines the process of OPTICS. In addition, Figure 2 shows the impact of different parameters set upon the results of OPTICS clustering (as compared to the employment of DBSCAN) with respect to the classical Iris dataset [27]. NOte that both algorithms share the same *MinSample* in defining the minimum size of a cluster, while DBSCAN purely depends on fixed epsilon value and OPTICS uses ξ to define reachability without a fixed epsilon value. It can be seen from

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 October 2023 doi:10.20944/preprints202310.1248.v1

the illustrations given in this figure that under different settings, results from the OPTICS clustering algorithm (regarding variation of ξ) exhibit more effective clusters (in smaller groups and located in more dense areas). This forms sharp contrast with the outcomes of applying DBSCAN where clusters give relatively larger clusters across varying densities, and it may even generate no clusters when encountering a greater *MinSample* value. The above empirical findings indicate that OPTICS offers better robustness while determining groups within a dense area of data points. In particular, this differs from the existing GA-based approach [22] that utilises pre-defined hypercubes to depict the knowledge space. Applying a clustering algorithm avoids strict boundaries like the hypercubes. Furthermore, the reachability feature in OPTICS introduces a new level of robustness on top of DBSCAN, which can determine tighter clusters, facilitating the approach proposed herein to investigate the generalisation of the interpolated rules with more informative content.

Algorithm 1 OPTICS Clustering Algorithm

Start:
Procedure OPTICS(Dataset D, MinPts, ξ, 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 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 reachability distance of n and add/update its position in $SeedList$.
  Step 4. Return $OrderedList$.
End.
Function RegionQuery($point$, ξ, maxSamples):
  Return all points within reachability distance considering ξ and restricted by maxSamples.
End.

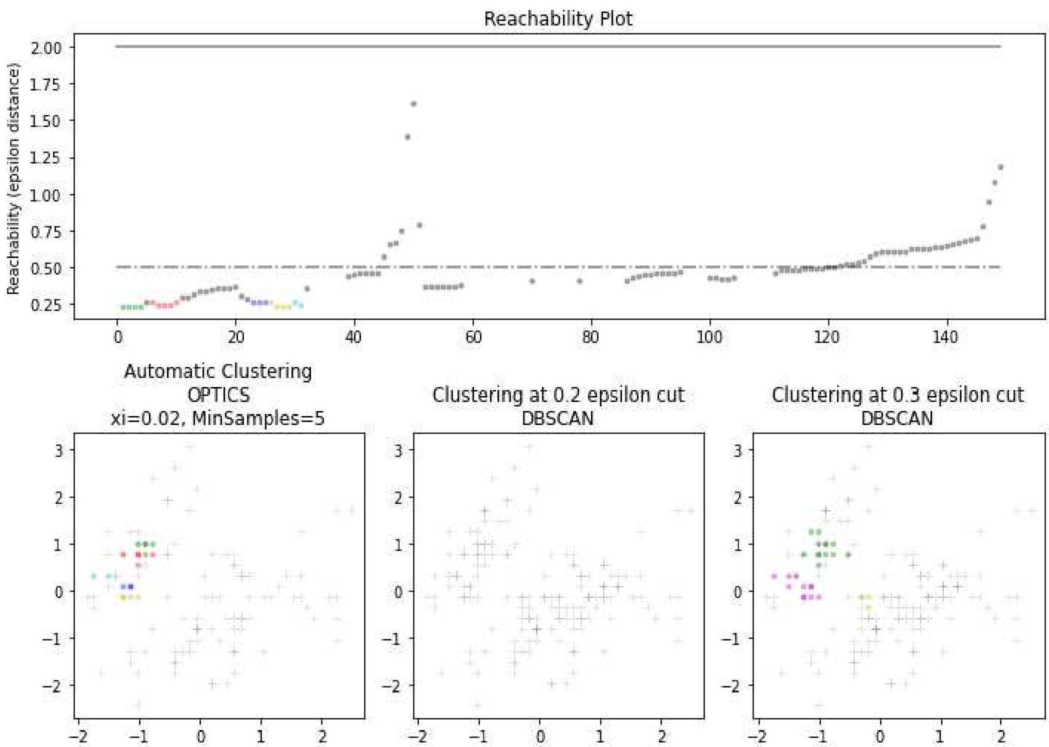

Figure 2. *Cont.*

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 October 2023 doi:10.20944/preprints202310.1248.v1

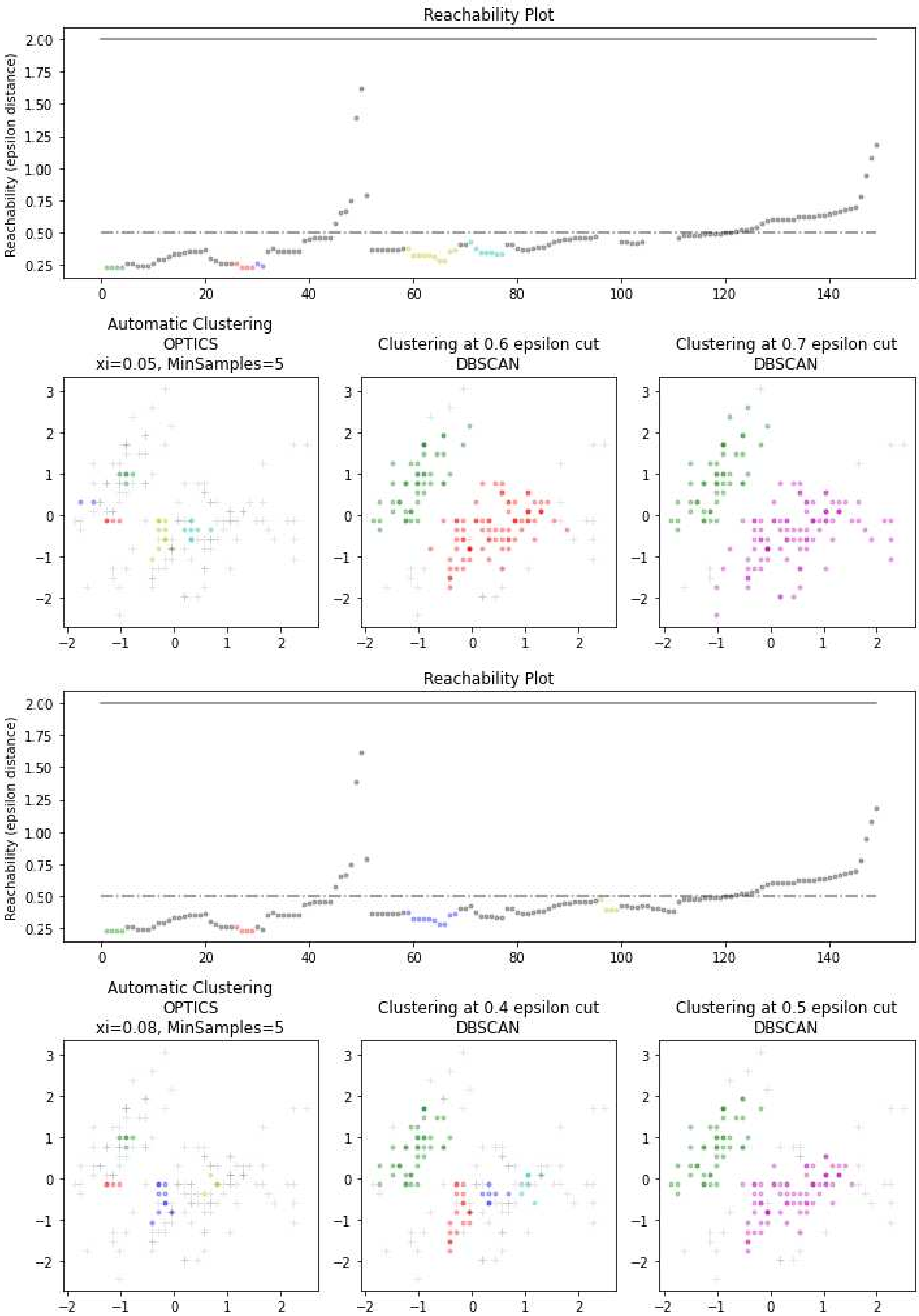

Figure 2. *Cont.*

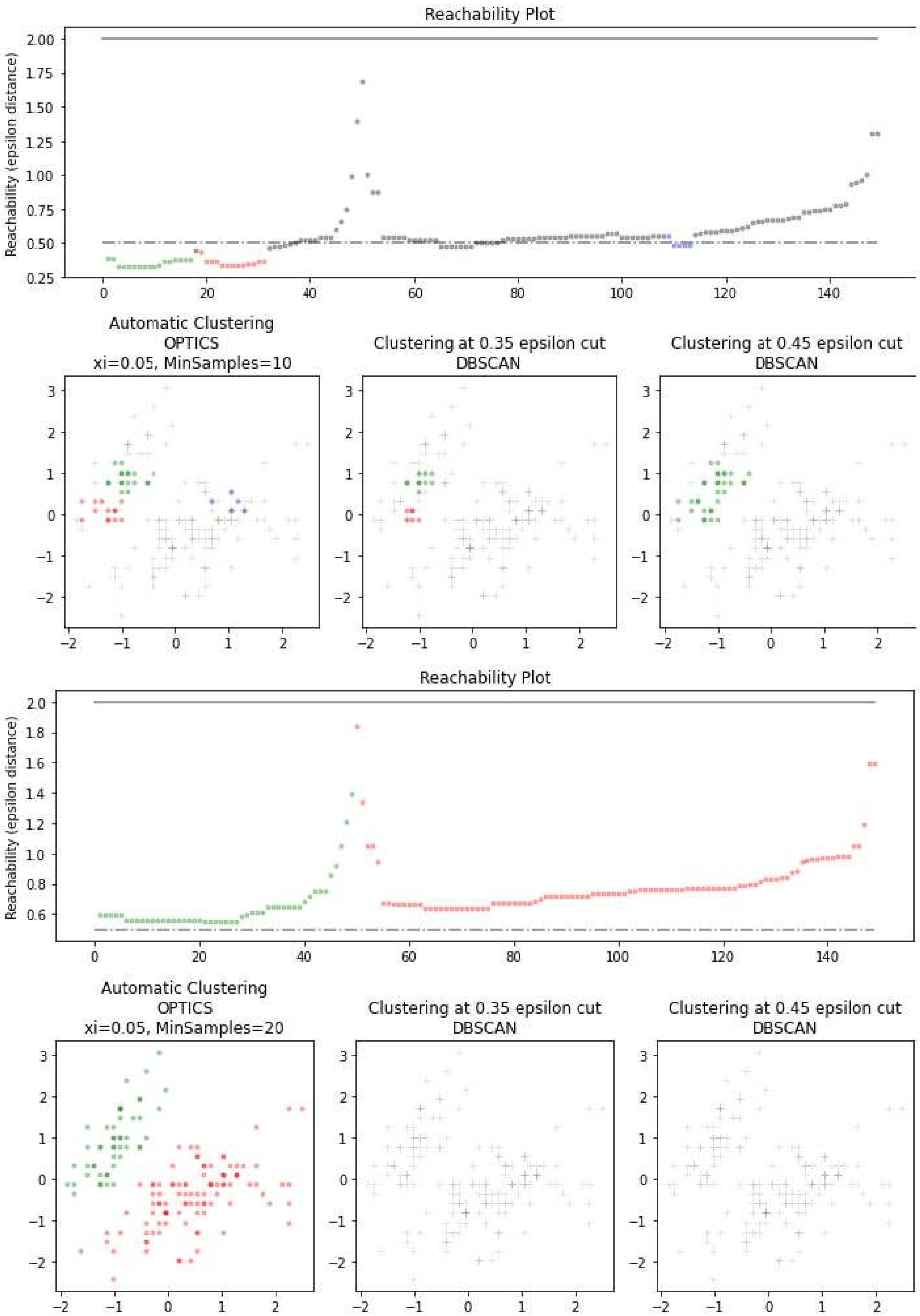

Figure 2. Different OPTICS clustering settings on Iris dataset.

3. Density Based Dynamic Fuzzy Rule Interpolation

This section describes the proposed method employing OPTICS to construct a dynamic fuzzy rule interpolation, hereafter abbreviated to OPTICS-D-FRI.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 October 2023 doi:10.20944/preprints202310.1248.v1

3.1. Structure of Proposed Approach

To provide an overview, Figure 3 illustrates the structure of OPTICS-D-FRI. Throughout the subsequent procedure, steps are enumerated with a notation '(k)', representing the sequence of operations as depicted in Figure 3.

Initiated with a sparse rule base, denoted as $\mathcal{R}$, and one of unseen observations (which has not been utilised to derive any outcome), 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 these 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).

Regarding every potential outcome B_z (8), the next stage computes a linguistic attribute frequency weight, as outlined in [26], 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 rules $\mathcal{R}_i^S : A_{m,1}, A_{m,2}, \ldots, A_{m,N}$ (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, the rule is selected (10) and promoted into the original sparse rule base, leading to an enriched or promoted rule base $\mathcal{P} = \mathcal{R} \cup \{\mathcal{R}_i^S\}$ (10). Conversely, if score S_{Bz} does not stand out within $Score_B$, the consequent B_z will be substituted by $B_z = B_{MAX(Score_B)}$, and then the entire process iterates.

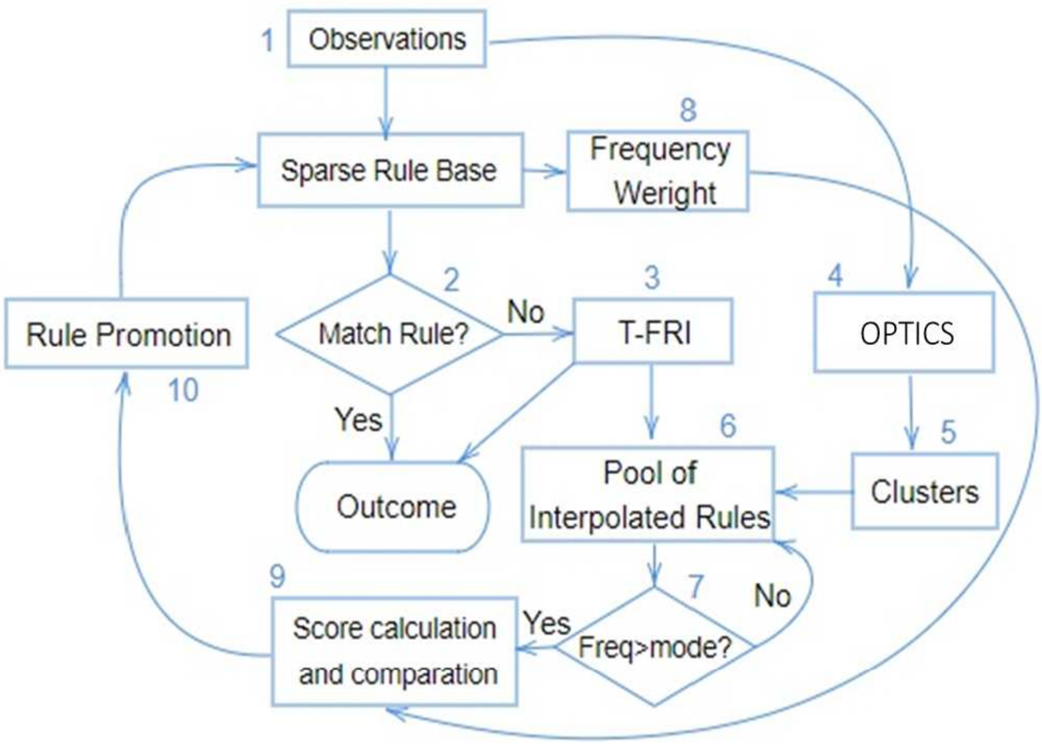

Figure 3. Proposed OPTICS-D-FRI framework.

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 $O_{OPTICS-D-FRI} = O(n^2)$.

3.2. Membership Frequency Weight Calculation and Interpolated Rule Promotion

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

Algorithm 2 Frequency Weight 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 consequent (where C is the total number of decision categories, e.g., 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 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 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 frequency weights table for each $Rep(B) \in \{0, 1, \cdots, C\}$.

Step 5. Sum 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 interpolated consequent with outcome of $Max(Score_B)$.

End.

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.

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.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 October 2023 doi:10.20944/preprints202310.1248.v1

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

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

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: The $Score_B = [S_{Rep(B)=0} : 1 + 0 + 1 = 2, S_{Rep(B)=1} : 0 + 1 + 0 = 1, S_{Rep(B)=2} : 0 + 0 + 1 = 1]$. In this case, $Rep(B) = 0$ is associated with the highest score, while the selected rule is concerned with the same outcome. Therefore, this selected interpolated rule is promoted, becoming a new rule in the sparse rule base.

4. Experimental Investigations

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

4.1. Datasets and Experimental Setup

As an initial attempt towards the development of a comprehensive dynamic fuzzy rule interpolative reasoning tool, five classical datasets, namely Iris, Yeast, Banana, Glass and Tae, 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 [28] in the R environment. Detailed description of all five datasets is presented in Table 4.

Table 4. Description of Datasets.

	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 other than the sizes of the partitioned sub-datasets being pre-specified: One of which 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 [29] facilitates the rule generation process. More advanced rule induction techniques (e.g., the method of [30]) may be utilised as the alternative if preferred, for enhanced outcomes. After the formation of a dense rule base $\mathcal{D}$, a strategic rule reduction by randomly dropping 30% of the rules from the dense rule base is done to create a sparse rule base, with the remaining 70% being regarded as the sparse rule base $\mathcal{R}$.

The other two sub-datasets, each containing 25% of the original data, are used separately for dynamic fuzzy rule interpolation (when no rules match a novel observation) and final testing. In particular, the sub-dataset for final testing is adopted to evaluate the performance of the interpolated and subsequently promoted rule base, $\mathcal{P}$. Additionally, by calculating the membership frequency weight of the promoted rule base $\mathcal{P}$ (which adds the promoted rule into the original sparse rule base and recalculates the frequency weights of the entire promoted rule base). As a result, the dynamically promoted rule base with weights $\mathcal{P}_{\mathcal{D}}$, denoted by OPTICS-D-FRI(DW), is also evaluated using this third sub-dataset.

A 10x10 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.

4.2. Experimental Results

The results of performance evaluation are presented in the following five tables, 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. 10 ×10 cross-validation on 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. 10×10 cross-validation on 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. 10×10 cross-validation on 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. 10×10 cross-validation on 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. 10×10 cross-validation on 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**

4.3. 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 antecedent space without any designated design, which may lead to the overall low accuracy on these datasets. This does not adversely affect this study since it is the relative performance that is considered amongst 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 in comparison to the use of traditional T-FRI.

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 can help improve the efficiency of the over 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 of 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.

5. Conclusion

This paper has proposed 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. It also introduces a novel way of determining populated regions, by clustering the new observations and annotating the group of interpolated rules, thereby eliminating the need for predefined hyper-cubes. Notably, the proposed OPTICS-D-FRI demonstrates enhanced performance over traditional FRI in inference accuracy.

Being an initial attempt to combine OPTICS and T-FRI for a dynamic FRI framework, this study also identifies several critical aspects for further refinement. For example, the current method requires specific predefined parameters, including $MinSample$ and ξ for OPTICS. Investigating a mechanism to automatically evaluate observations and determine such parameters remains as further research. Also, 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.

Recent advancements in the FRI literature highlight the efficacy of approximate reasoning that utilise attribute-weighted rules or rule base stricture [31], with an emphasis on employing attribute ranking mechanisms, as proposed in [32,33]. A density-based approach for fuzzy rule interpolation has also been proposed recently [34]. Frequency weighting benefits both attribute ranking and FRI, especially when integrating higher-order FRI [35] with rule induction and feature selection strategies [36], to address more complex application scenarios. Another exciting exploration involves assessing dynamically weighted attribute frequency calculations, particularly with the inclusion of newly promoted rules. Lastly, while the present validation is based on five classical datasets, broader empirical investigations involving more complex datasets, such as mammographic risk analysis [37,38], are necessary to reveal the potential and limitations of the proposed framework fully.

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.; writing—original draft preparation, J.L..; writing—review and editing, C.S. and Q.S.; visualization, J.L.; supervision, C.S. and Q.S.; project administration, J.L. and 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 earlier conference paper, which was recommended for consideration of producing an expanded version for inclusion in this special issue. The first author is supported in part by Aberystwyth University PhD scholarships.

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

Abbreviations

Abbreviations

The following abbreviations are used in this manuscript:

MDPI	Multidisciplinary Digital Publishing Institute
DOAJ	Directory of Open Access Journals
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

References

1. Dubois, D.; Prade, H. On fuzzy interpolation. *International Journal Of General System* **1999**, *28*, 103–114.
2. Kóczy, L.; Hirota, K. Approximate reasoning by linear rule interpolation and general approximation. *International Journal of Approximate Reasoning* **1993**, *9*, 197–225.
3. Kóczy, L.; Hirota, K. Interpolative reasoning with insufficient evidence in sparse fuzzy rule bases. *Information Sciences* **1993**, *71*, 169–201.
4. Saga, S.; Makino, H.; Sasaki, J.I. A method for modeling freehand curves—The fuzzy spline interpolation. *Systems and Computers in Japan* **1995**, *26*, 77–87.
5. Li, F.; Shang, C.; Li, Y.; Yang, J.; Shen, Q. Approximate reasoning with fuzzy rule interpolation: background and recent advances. *Artificial Intelligence Review* **2021**, *54*, 4543–4590.
6. Gedeon, T.; Kóczy, L. Conservation of fuzziness in rule interpolation. Proc. of the Symp. on New Trends in Control of Large Scale Systems, 1996, Vol. 1, pp. 13–19.
7. Hsiao, W.H.; Chen, S.M.; Lee, C.H. A new interpolative reasoning method in sparse rule-based systems. *Fuzzy Sets and Systems* **1998**, *93*, 17–22.
8. Kóczy, L.T.; Hirota, K.; Gedeon, T.D. Fuzzy Rule Interpolation by the Conservation of Relative Fuzziness. *J. Adv. Comput. Intell. Intell. Informatics* **2000**, *4*, 95–101.
9. Baranyi, P.; Gedeon, T.D.; Koczy, L. Rule interpolation by spatial geometric representation. Proc. IPMU, 1996, Vol. 96, pp. 483–488.
10. Bouchon-Meunier, B.; Marsala, C.; Rifqi, M. Interpolative reasoning based on graduality. Ninth IEEE International Conference on Fuzzy Systems. FUZZ-IEEE 2000 (Cat. No. 00CH37063). IEEE, 2000, Vol. 1, pp. 483–487.
11. Huang, Z.; Shen, Q. Fuzzy interpolative reasoning via scale and move transformations. *IEEE Transactions on Fuzzy Systems* **2006**, *14*, 340–359.
12. Huang, Z.; Shen, Q. Fuzzy interpolation and extrapolation: A practical approach. *IEEE Transactions on Fuzzy Systems* **2008**, *16*, 13–28.
13. Jenei, S.; Klement, E.P.; Konzel, R. Interpolation and extrapolation of fuzzy quantities–The multiple-dimensional case. *Soft Computing* **2002**, *6*, 258–270.
14. Johanyák, Z.C.; Kovács, S. Fuzzy rule interpolation by the least squares method. 7th International Symposium of Hungarian Researchers on Computational Intelligence (HUCI 2006), 2006, Vol. 2425, pp. 495–506.
15. Wu, Z.Q.; Masaharu, M.; Shi, Y. An improvement to Kóczy and Hirota's interpolative reasoning in sparse fuzzy rule bases. *International Journal of Approximate Reasoning* **1996**, *15*, 185–201.

16. Lee, L.W.; Chen, S.M. Fuzzy interpolative reasoning for sparse fuzzy rule-based systems based on the ranking values of fuzzy sets. *Expert systems with applications* **2008**, *35*, 850–864.
17. Åström, K.J.; Wittenmark, B. *Adaptive control*; Courier Corporation, 2013.
18. Liu, B.D.; Chen, C.Y.; Tsao, J.Y. Design of adaptive fuzzy logic controller based on linguistic-hedge concepts and genetic algorithms. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* **2001**, *31*, 32–53.
19. Mohan, S.; Bhanot, S. Comparative study of some adaptive fuzzy algorithms for manipulator control. *International Journal of Computational Intelligence* **2006**, *3*, 303–311.
20. Wang, Y.; Deng, H.; Chen, Z. Adaptive fuzzy logic controller with rule-based changeable universe of discourse for a nonlinear MIMO system. 5th International Conference on Intelligent Systems Design and Applications (ISDA'05). IEEE, 2005, pp. 8–13.
21. Labiod, S.; Boucherit, M.S.; Guerra, T.M. Adaptive fuzzy control of a class of MIMO nonlinear systems. *Fuzzy sets and systems* **2005**, *151*, 59–77.
22. Naik, N.; Diao, R.; Shen, Q. Dynamic fuzzy rule interpolation and its application to intrusion detection. *IEEE Transactions on Fuzzy Systems* **2017**, *26*, 1878–1892.
23. Ester, M.; Kriegel, H.P.; Sander, J.; Xu, X.; others. A density-based algorithm for discovering clusters in large spatial databases with noise. kdd, 1996, Vol. 96, pp. 226–231.
24. Ankerst, M.; Breunig, M.M.; Kriegel, H.P.; Sander, J. OPTICS: Ordering points to identify the clustering structure. *ACM Sigmod record* **1999**, *28*, 49–60.
25. Lin, J.; Shang, C.; Shen, Q. Towards Dynamic Fuzzy Rule Interpolation with Harmony Search. Proceedings of the 21st Annual Workshop on Computational Intelligence., 2022.
26. Lin, J.; Xu, R.; Shang, C.; Shen, Q. Towards Dynamic Fuzzy Rule Interpolation via Density-Based Spatial Clustering of Interpolated Outcomes. Proceedings of the 2023 IEEE International Conference on Fuzzy Systems. IEEE Press, 2023.
27. Fisher, R.A. The use of multiple measurements in taxonomic problems. *Annals of eugenics* **1936**, *7*, 179–188.
28. DeBruine, L. *faux: Simulation for Factorial Designs*, 2021. R package version 1.1.0, doi:10.5281/zenodo.2669586.
29. Wang, L.X.; Mendel, J. Generating fuzzy rules by learning from examples. *IEEE Transactions on Systems, Man, and Cybernetics* **1992**, *22*, 1414–1427.
30. Chen, T.; Shang, C.; Su, P.; Shen, Q. Induction of accurate and interpretable fuzzy rules from preliminary crisp representation. *Knowledge-Based Systems* **2018**, *146*, 152–166.
31. Jiang, C.; Jin, S.; Shang, C.; Shen, Q. Towards utilization of rule base structure to support fuzzy rule interpolation. *Expert Systems* **2023**, *40*, e13097.
32. Li, F.; Li, Y.; Shang, C.; Shen, Q. Fuzzy knowledge-based prediction through weighted rule interpolation. *IEEE Transactions on Cybernetics* **2020**, *50*, 4508–4517.
33. Li, F.; Shang, C.; Li, Y.; Yang, J.; Shen, Q. Interpolation with just two nearest neighbouring weighted fuzzy rules. *IEEE Transactions on Fuzzy Systems* **2020**, *28*, 2255–2262.
34. Gao, F. Density-based approach for fuzzy rule interpolation. *Applied Soft Computing* **2023**, *143*, 110402.
35. Chen, C.; MacParthaláin, N.; Li, Y.; Price, C.; Quek, C.; Shen, Q. Rough-fuzzy rule interpolation. *Information Sciences* **2016**, *351*, 1–17.
36. Jensen, R.; Cornelis, C.; Shen, Q. Hybrid fuzzy-rough rule induction and feature selection. 2009 IEEE International Conference on Fuzzy Systems, 2009, pp. 1151–1156.
37. Li, F.; Shan, C.; Li, Y.; Shen, Q. Interpretable mammographic mass classification with fuzzy interpolative reasoning. *Knowledge-Based Systems* **2020**, *191*, 105279.
38. Dounis, A.; Avramopoulos, A.N.; Kallergi, M. Advanced Fuzzy Sets and Genetic Algorithm Optimizer for Mammographic Image Enhancement. *Electronics* **2023**, *12*, 3269.