# Explainability and Efficiency in Fuzzy Logic Systems: Frameworks for Temporal Classification, Monotonicity, and Hybrid Predictive Modeling

**By**

**Ashish Bhatia**

**A thesis submitted for the degree of**

**Doctor of Philosophy**

**School of Computer Science and Electronic Engineering**

**University of Essex**

**April 2026**

# Abstract

This thesis advances fuzzy logic systems by improving explainability without compromising predictive accuracy in regression and time-series classification tasks. Motivated by the persistent trade-off between interpretability and performance in fuzzy methodologies, particularly in regulated domains requiring transparent decisions, the research develops novel frameworks addressing temporal modelling, semantic consistency, and hybrid inference. The work demonstrates that fuzzy systems can achieve semantic alignment with domain knowledge without compromising effectiveness. Key contributions include:

The Optimized Time Series Interval Type-2 Fuzzy System framework introduces correlation-based time step selection with polynomial fitting for automated identification of predictive temporal patterns. This enables optimised fuzzy set generation over time periods, creating linguistic features (e.g., ”recent” or ”long-term”) that enhance interpretability and noise robustness. Empirical evaluations on UCR classification benchmarks indicate clear advancements over baseline results, while providing explainable rules closely aligned with human temporal understanding.

For semantic inconsistencies in fuzzy regression, the thesis formalises logical gaps, violations of monotonic relationships, and proposes a detection framework with quality metric guided rule insertion. Validated on a credit scoring dataset, this repair mechanism restores consistency via targeted interventions, preserving transparency with minimal impact on predictive accuracy and rule base size.

In hybrid modelling, a novel Mamdani-TSK framework integrates linguistic consequents with constrained polynomials and dual weights optimised via ant colony algorithms. Benchmark results on KEEL datasets demonstrate RMSE improvements of up to 19% over fuzzy baselines, achieving superior precision whilst maintaining full explainability.

These innovations collectively extend fuzzy logic’s theoretical foundations, offering reusable methodologies for explainable AI across classification and regression. The research pro-

vides empirical evidence from multiple domains, highlighting the practical value of these approaches in high-stakes applications. This work reinforces fuzzy logic's role in responsible AI, showing that enhanced explainability through systematic innovations can yield trustworthy systems in an increasingly regulated world.

# Acknowledgments

I would like to express my deepest gratitude to all those who contributed to the completion of this work. There were unique challenges in being a part-time student while managing a full-time job and this academic journey wouldn't have been possible without your support, guidance, and encouragement.

First and foremost, I extend my sincere appreciation to my initial supervisor, Prof. Hani Hagras, for his invaluable guidance. His support was instrumental in shaping this work. I am also very grateful to my supervisors towards the later part of this research, Dr. Renato Cordeiro de Amorim and Dr. Vito De Feo, for their support and guidance in getting this work over the finishing line.

Last but not least, my heartfelt gratitude goes to my wife Devayani, and my son Aniket. Devayani has, at times, believed in my ability to complete this work more than I believed it myself. Aniket was born during my studentship and now at 6 years old, is not fully sure why his father still needs to continue studying but extends his unwavering support nonetheless. Without the support of these two, this work might have been completed much sooner but wouldn't have been as much fun and rewarding.

# List of Publications

## Conference

- A. Bhatia and H. Hagras, ”A time series based explainable interval type-2 fuzzy logic system,” in 2022 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), 2022, pp. 1-7.
- A. Bhatia and H. Hagras, ”A Type-2 Fuzzy Time Series Classification System with Optimized Time Period Selection,” in 2024 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), 2024, pp. 1-9.
- A. Bhatia and H. Hagras, ”Identifying and Rectifying Rational Gaps in Fuzzy Rule Based Systems for Regression Problems,” in 2021 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), 2021, pp. 1-6.

# Contents

# List of Figures

# Chapter 1

# Introduction

## 1.1 Introduction

The field of artificial intelligence (AI) and machine learning (ML) has witnessed unprecedented growth, driven by advancements in computational power and data availability. Fuzzy logic systems have emerged as a powerful paradigm for handling uncertainty, imprecision, and human-like reasoning in predictive modelling tasks [1]. Fuzzy Rule-Based Systems (FRBSs) offer a unique blend of interpretability and flexibility, making them particularly suitable for applications requiring explainable decisions, such as healthcare diagnostics, financial forecasting, and industrial control systems [2].

However, traditional fuzzy systems often struggle with the trade-off between accuracy and explainability, especially in complex tasks like time-series prediction and regression modelling. This thesis explores three interconnected research areas within fuzzy logic:

1. **Enhancing explainability in fuzzy time-series approaches** – addressing the limitations of traditional methods while enhancing interpretability and explainability.

2. **Addressing logical gaps in fuzzy regression systems** – ensuring semantic consistency in rule-based predictions.

3. **Developing hybrid fuzzy regression models** – achieving improved accuracy while preserving interpretability and explainability.

### 1.1.1 Time-Series Classification Challenge

#### Introduction

Time-series analysis has emerged as a cornerstone of modern data analytics, with applications spanning financial market forecasting, healthcare monitoring, climate modelling, and industrial process control [3], [4]. The ability to accurately forecast future values or classify temporal patterns enables critical decision-making across these domains [5], [6]. As temporal data volume and complexity continue to grow exponentially, developing sophisticated yet interpretable analysis methods has become increasingly crucial.

Recent advances in machine learning have led to remarkable improvements in time-series forecasting and classification accuracy. Deep learning models such as Long Short-Term Memory (LSTM) networks and Transformer-based architectures [7] [8], along with ensemble approaches like HIVE-COTE (Hierarchical Vote Collective of Transformation Ensembles) [9], achieve state-of-the-art performance by learning complex temporal dependencies through multiple layers of non-linear transformations, capturing long-range interactions via attention mechanisms, and extracting hierarchical features without manual engineering [10] [11].

However, these impressive capabilities come with significant limitations. Deep learning models require large training datasets [12] and substantial computational resources [13] [14], restricting their applicability in data-scarce scenarios such as rare disease prediction, newly emerging market analysis and medical applications, where they often suffer from severe over-fitting [15, 16]. More critically, the accuracy gains have come at the cost of interpretability. These models operate as opaque boxes—their defining characteristics of multiple non-linear transformation layers, millions of parameters, and complex interaction effects render them incomprehensible to human understanding [17] [18]. This opacity poses substantial challenges in domains where understanding the rationale behind predictions is as important as the predictions themselves, such as medical diagnosis and financial decision-making [6, 19, 20].

Furthermore, deep learning models struggle to incorporate domain knowledge effectively. While traditional statistical methods and fuzzy systems can integrate expert knowledge through prior distributions or rule specifications, deep learning typically learns purely from data, potentially missing important domain constraints or physical laws [21, 22]. This can lead to predictions that, while statistically accurate on test sets, violate fundamental domain principles [23, 24].

The demand for explainable models has intensified due to both regulatory requirements and practical necessities [25]. The European Union's GDPR explicitly mandates a "right to explanation" for automated decision-making systems [26]. Beyond compliance, explainable models enable domain experts to validate predictions, identify biases, and build trust in automated systems—critical factors in medical diagnosis, financial risk assessment, and infrastructure management [27, 28].

Among explainable approaches, fuzzy logic systems offer unique advantages for time-series analysis. Fuzzy Rule-Based Systems (FRBS) naturally align with human cognitive processes through linguistic variables and IF-THEN rules that mirror expert reasoning [1,29], enabling stakeholders to understand model behaviour without deep technical expertise. They also excel at handling the inherent uncertainty and imprecision in real-world temporal data [30, 31].

Several fuzzy approaches to time-series analysis have been developed. Fuzzy Time Series (FTS), introduced by Song and Chissom [32], represents time series as sequences of fuzzy sets and mines Fuzzy Logical Relations (FLR) between consecutive periods. While conceptually elegant for forecasting, this approach restricts each observation to a single fuzzy set, contradicting the fundamental principle of partial membership in multiple sets. These approaches are discussed in depth in Section 2.3.

For classification tasks, alternative approaches include treating historical time-steps as features in Fuzzy Inference Systems or employing fuzzy clustering techniques [33]. Other methods encompass fuzzy cognitive maps, which model causal relationships but suffer from computational intensity [34]; fuzzy rule mining, effective for pattern discovery yet prone to rule explosion [35]; and F-transform-based techniques, which provide noise reduction but may lose fine-grained details [36]. While modern extensions—such as recurrence-aware Long-Term Cognitive Networks (LTCNs)—address issues like convergence and computational intensity [37], these enhancements primarily improve performance without advancing inherent explainability, particularly in models with large feature spaces.

Despite their theoretical advantages, existing fuzzy and traditional explainable methods face critical limitations in time-series classification:

- **Feature explosion in high-dimensional temporal data**: The number of fuzzy rules grows rapidly with increasing dimensionality and series length, leading to computational inefficiency and scalability challenges.
- **Noise sensitivity in real-world applications**: Temporal noise, irregular sampling, or delays can shift pattern occurrences in time, degrading model performance.
- **Limited temporal abstraction capabilities**: Traditional fuzzy approaches typically model only short-term dependencies and lack mechanisms for capturing long-range temporal patterns.

These limitations are further explored in Section 3.2. The current state of knowledge reveals a gap in methods that provide linguistic temporal representations aligned with human reasoning while maintaining competitive performance—a gap this research aims to address.

### 1.1.2 Logical Gaps in Fuzzy Regression

The second area examines logical gaps in fuzzy regression systems. Machine learning has become a core component of contemporary decision-making systems across a wide range of application domains. As these systems increasingly influence high-stakes outcomes, attention has shifted from raw predictive accuracy toward the internal soundness of the models themselves. In particular, there is growing concern about whether model outputs are supported by logically coherent internal structures, rather than being the product of fragmented or weakly connected reasoning processes.

A central challenge lies in the presence of *logical gaps*: regions where a model's internal rules, representations, or decision pathways fail to adequately support its predictions. Such gaps may arise when parts of the input space are insufficiently covered, when relationships between variables are internally inconsistent, or when learned patterns contradict basic domain knowledge. These issues can remain hidden in highly opaque models, but they become especially salient in systems whose internal mechanisms are explicitly exposed.

Models designed for transparency are often favored in regulated or safety-critical settings because they allow human inspection of their reasoning process. However, transparency alone does not ensure logical completeness or consistency. A model may be fully interpretable yet still exhibit gaps between its stated rules and the conclusions it produces. When this occurs, users can clearly observe not only the model's decisions but also the absence of sufficient justification for those decisions, undermining confidence and usability.

Fuzzy rule-based systems (FRBSs) represent a prominent class of models intended to balance interpretability with expressive power. Since their introduction through Zadeh's fuzzy logic framework, these systems have encoded knowledge using linguistic IF–THEN rules, enabling analysts to examine individual components of the reasoning process. In regression settings, common architectures such as Mamdani and Takagi–Sugeno–Kang models combine multiple local rules to approximate complex input–output relationships. Their rule-based structure makes it possible to inspect coverage of the input space, identify redundancies, and trace outputs back to human-readable statements.

Despite these advantages, FRBSs are not immune to logical gaps. In many domains, well-established qualitative relationships are expected to hold across the entire input space. For example, increases in a risk-related variable should not lead to decreases in predicted harm, and increases in a causal input should not produce counterintuitive reductions in response. When a rule-based model violates such expectations, the issue is not a lack of interpretability but a failure of internal logical coherence.

Data-driven construction methods can exacerbate this problem. To achieve compactness and reduce complexity, rule induction algorithms often prune candidate rules aggressively. While this improves parsimony, it can also leave portions of the input space weakly supported or entirely uncovered, creating discontinuities or unjustified transitions between rules. The resulting model may appear clear and interpretable, yet still rely on incomplete or logically disconnected reasoning structures.

Logical gaps impact a number of industries and use cases where regulatory oversight and transparency are critical factors for model based decision making. In particular, financial services illustrate both the importance and the cost of logical coherence. FRBSs are used for credit approval, portfolio monitoring, bankruptcy prediction, and financial security assessment [38–40].

Understanding, detecting, and mitigating logical gaps is therefore essential for ensuring that transparent models are not only readable but also logically robust. Addressing this issue is a prerequisite for deploying interpretable systems that can be trusted to behave consistently with established domain principles across all relevant operating conditions.

### 1.1.3 Fuzzy Regression Modelling Trade-offs

Regression analysis is a fundamental method in statistical modelling and machine learning, used to examine and quantify the relationships between variables (inputs) and a continuous dependent variable (output). This technique serves as a critical tool for predictive tasks, facilitating the analysis of patterns and trends across various fields. Its applications span numerous domains, including finance, healthcare, engineering, and environmental science, where it has driven significant advancements in prediction, forecasting, and diagnostic capabilities [41]. In regulated sectors such as banking, healthcare, and insurance, regression models must not only achieve high accuracy but also provide transparent decision-making processes to comply with regulatory requirements [42]. The European Union's General Data Protection Regulation (GDPR) explicitly grants individuals the right to meaningful explanations of algorithmic decisions, making model interpretability a legal requirement rather than merely a desirable feature [43].

Traditional regression techniques, such as linear regression and support vector regression, have proven effective in handling numerous problems [44]. Linear regression provides mathematical transparency through interpretable coefficients and robust theoretical foundations, offering clear insights into variable relationships [45]. This makes it particularly valuable in regulated environments, where decision rationales must be documented and audited [18]. However, these methods often encounter limitations when dealing with the complexities of real-world data. The presence of uncertainty, ambiguity, and linguistic information in datasets poses challenges for traditional approaches that assume precise and well-defined numerical relationships [46]. Furthermore, their linear assumptions frequently fail to capture the intricate non-linear patterns present in modern datasets, limiting their predictive accuracy in complex domains [47]. Linear models may not be ideal for cases like financial time series forecasting, where non-linear volatility clustering is common [48], biological dose-response relationships with sigmoidal curves [49], or image recognition tasks, where pixel relationships are inherently non-linear [50].

In recent years, advancements in neural networks and deep learning have driven significant progress in various machine learning tasks, including regression. Deep learning models, with their ability to learn complex non-linear relationships from vast amounts of data, have achieved state-of-the-art results in many domains [12]. However, deep learning models lack interpretability, making it difficult to understand the underlying reasoning behind their predictions. This "black-box" nature creates significant barriers in regulated industries where model decisions must be explainable, auditable, and justifiable to stakeholders and regulators [17]. Financial institutions face scrutiny from regulators when using opaque models for credit scoring or risk assessment [51]. Healthcare providers must justify diagnostic recommendations to patients and insurance companies [52]. The opacity of deep learning models can lead to regulatory non-compliance, legal challenges, and erosion of public trust [28, 53]. In addition, deep learning models, characterised by their large number of parameters, are generally ill-suited for training on small datasets due to the risk of over-fitting and insufficient data representation [54].

The tension between accuracy and interpretability has become a critical challenge in machine learning research and practice [17]. Recent research has shown that in many regulated sectors, slightly less accurate but fully interpretable models are preferred over marginally more accurate black-box alternatives [17]. This preference stems from legal requirements, ethical considerations, and the need for human oversight in critical decisions. Recent years have brought increased attention to explainable and interpretable model research, yet challenges remain in model innovations striving to balance accuracy and explainability [55, 56].

As discussed in Section 2.2, fuzzy logic offers a framework for representing and reasoning with uncertain or imprecise information [57]. By incorporating human-like reasoning and linguistic rules, fuzzy logic systems provide a powerful mechanism for handling problems characterised by vagueness and subjectivity. The interpretable nature of fuzzy rules aligns well with regulatory requirements for explainable models [58]. In regression, fuzzy approaches have gained prominence due to their ability to model data relationships that are not easily captured by traditional methods [59]. Fuzzy systems have demonstrated particular value in regulated sectors where combining expert knowledge with data-driven insights is essential [60].

The Mamdani and Takagi-Sugeno-Kang (TSK) approaches are the two main fuzzy system paradigms applied in regression [29, 61]. Mamdani systems are known for their interpretability and transparency, making them particularly suitable for regulated environments [62]. However, enhancing interpretability often requires limiting linguistic terms, resulting in coarse granularity that reduces regression accuracy [2, 63].

TSK systems improve numerical precision by employing crisp mathematical functions in the consequents and can achieve performance comparable to neural networks [64, 65]. Nevertheless, these mathematical consequents reduce linguistic interpretability and obscure the semantic meaning of rules [63, 66], which is problematic in regulated contexts requiring natural-language explanations.

The absence of methodologies that achieve both high accuracy and meaningful interpretability limits adoption in domains such as healthcare and finance, where both performance and explainability are legally and ethically required [67, 68].

The literature highlights a persistent challenge in balancing these attributes, particularly in regulated sectors demanding both high performance and auditable decisions [17].

**These areas, while distinct, share an overarching research gap: the need for fuzzy systems that achieve high accuracy without sacrificing genuine interpretability. Studying them together reveals common themes in handling uncertainty through interval type-2 fuzzy sets and evolutionary optimisation, advancing the broader goal of explainable AI in predictive modelling [56].**

## 1.2 Problem Statement and Research Rationale

The primary research problems addressed in this thesis are: **(a)** the persistent tension between accuracy and explainability, and **(b)** the enhancement of explainability in fuzzy logic systems for predictive tasks. These challenges manifest differently across tasks:

- **Time-series classification**: Complex ensemble approaches dominate performance but lack transparency [9], while fuzzy approaches offer interpretability but sometimes at the expense reduced accuracy.
- **Regression modelling**: Fuzzy systems may produce semantically inconsistent predictions that violate domain knowledge, or sacrifice explainability for precision.

The rationale for investigating these areas stems from:

1. **Regulatory requirements**: The requirement for enhanced explainability is particularly strong in regulated industries where models must be both accurate and justifiable [67]. Regulations like the EU's GDPR mandate "right to explanation" for automated decisions [43], and

2. **Domain-specific constraints**: Critical applications require both high performance and interpretable decisions.

This research addresses the objective of developing trustworthy fuzzy systems by simultaneously addressing three key challenges:

- Logical gaps in fuzzy regression,
- Temporal explainability in fuzzy time series classification, and
- Improving accuracy in fuzzy regression systems without losing explainability.

The research enhances both explainability and accuracy utilising techniques such as interval type-2 fuzzy sets and evolutionary algorithms.

## 1.3 Research Objectives

The primary research objective is to develop fuzzy logic methodologies that enhance explainability while improving or maintaining accuracy and semantic consistency in predictive tasks.

### 1.3.1 Fuzzy Time-Series Approaches

**Objectives:**

1. Design an automated method for temporal feature extraction that aligns with human reasoning

2. Develop an optimised interval type-2 fuzzy system using the new temporal features for time-series classification

3. Evaluate the trade-off between explainability and performance on benchmark datasets

### 1.3.2 Logical Gaps in Fuzzy Regression

**Objectives:**

1. Formalise the concept of logical gaps in FRBSs
2. Develop algorithms for gap detection and repair while preserving model transparency
3. Validate the approach on real-world datasets with monotonic relationships

### 1.3.3 Mamdani-TSK Hybrid Fuzzy Regression

**Objectives:**

1. Create a rule structure combining Mamdani and TSK elements
2. Develop an optimised interval type-2 fuzzy system using using the new rule structure
3. Compare hybrid models against pure fuzzy and non-fuzzy benchmarks

These area-specific questions contribute to the main objective by addressing complementary aspects of explainability:

- **Time-based linguistic features** in time-series classification
- **Semantic consistency** in regression
- **Precision-interpretability balance** in regression

Together, these objectives advance fuzzy systems toward algorithms that are transparent, semantically consistent, and accurate—essential qualities for domains where explainability is a mandatory requirement for predictive models.

## 1.4 Research Contributions

### 1.4.1 Time-Series

The time-series chapter introduces OTS-IVFS, which introduces novel techniques for linguistic temporal feature generation, automated time period discovery, and noise-robust aggregation mechanisms. These advances significantly enhance explainable time-series classification capabilities by providing both superior performance and interpretable temporal patterns.

### 1.4.2 Logical Gaps

The logical gaps chapter provides a novel logical gap detection algorithm for semantic inconsistencies alongside repair mechanisms that restore monotonicity with minimal accuracy loss. The research demonstrates that small rule insertions can close substantial gaps while preserving model transparency throughout the repair process. This approach ensures that fuzzy systems maintain their interpretability even as logical consistency is enforced.

### 1.4.3 Hybrid Regression

The hybrid regression chapter delivers a novel Hybrid Interval Type-2 Mamdani-TSK (HIT2-MTSK) approach with constrained polynomials. The research introduces a new rule type with Mamdani and TSK consequents and a dual weight mechanism. The framework enables precision without interpretability loss, providing a principled approach for balancing competing objectives in fuzzy modelling. HIT2-MTSK represents a significant advance in combining the strengths of different fuzzy modelling approaches.

These contributions range from temporal explainability to semantic consistency and precision enhancement, collectively advancing fuzzy systems towards higher-accuracy, interpretable models.

## 1.5 Thesis Structure

The rest of the thesis is organised as follows:

- **Chapter 2**: Reviews fuzzy logic foundations establishing the theoretical basis required for the methodological developments that follow.

- **Chapter 3**: Details time-series enhancements, discussing the time-series research gaps and the related research objectives identified in Chapter 1 and 2 in more depth, presenting the proposed methodology and experimental results.

- **Chapter 4**: Addresses logical gaps, discussing the issues arising from semantic gaps in fuzzy rule bases, formalizing the definition of logical gaps and describing detection/repair algorithms and real-world validation.

- **Chapter 5**: Discusses the need for improvements to the current fuzzy regression approach, introduces the hybrid regression model, including implementation details and comprehensive benchmarks.
- **Chapter 6**: Concludes by synthesising findings from all components, highlighting contributions and outlining future research directions derived from the presented results.

# Chapter 2

# Background

## 2.1 Introduction

This chapter introduces concepts that are used throughout the remaining chapters. Section 2.2 presents the background on Fuzzy Rule Based Systems. Section 2.3 gives an oveview of popular fuzzy time series prediction approaches followed by a brief review of classification and regression metrices in Section 2.4. These have been used in evaluation of results from the experiments.

## 2.2 Fuzzy Rule Based Systems

### 2.2.1 Introduction

Fuzzy Rule-Based Systems (FRBSs) are rule-based systems that employ linguistic fuzzy variables as antecedents and consequents to represent human-understandable knowledge [69]. These systems have been extensively applied across various domains in soft computing literature for over 40 years [69]. FRBSs effectively handle uncertainty and imprecision inherent in knowledge representation, making them more robust compared to classical logic systems [1].

### 2.2.2 Classical Mamdani Fuzzy Rule-Based Systems

The classical fuzzy systems based on Mamdani's approach [70] constitute the foundation of FRBSs. The generic structure comprises several key components as illustrated in Figure 2.1.

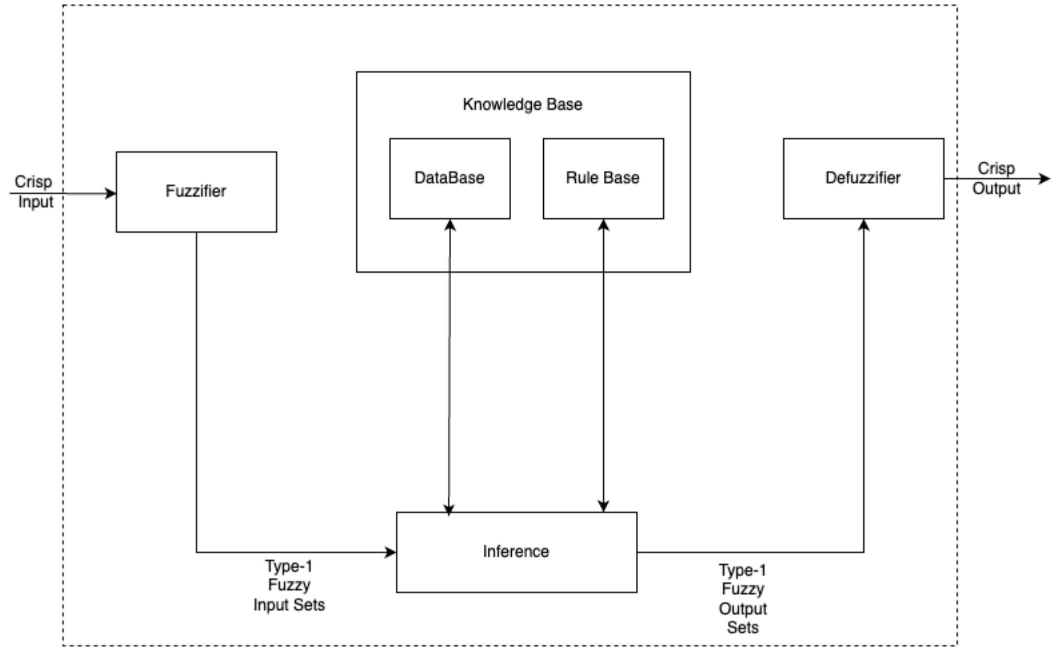


Figure 2.1: Structure of a Fuzzy Rule Based System

### Architecture and Components

A Mamdani FRBS consists of the following modules, each performing specific transformations on the input data for the regression or classification system.

- **Fuzzification Module**: Establishes a mapping between real-valued input data to fuzzy values based on membership functions. For a crisp input value $x_0$, the degree of membership in fuzzy set $A$ is calculated as:

$$\mu_A(x_0) = f_A(x_0)$$

where $f_A$ is the membership function for fuzzy set $A$.

Some common type-1 membership functions are shown in Figure 2.2. These are:

– *Triangular*

$$\mu_A(x) = \begin{cases} 0 & x \leq a \\ \frac{x-a}{b-a} & a < x \leq b \\ \frac{c-x}{c-b} & b < x < c \\ 0 & x \geq c \end{cases}$$

where the points $a$, $b$ and $c$ are as shown in Figure 2.2.

– *Trapezoidal*

$$\mu_A(x) = \begin{cases} 0 & x \leq a \\ \frac{x-a}{b-a} & a < x \leq b \\ 1 & b < x \leq c \\ \frac{d-x}{d-c} & c < x < d \\ 0 & x \geq d \end{cases}$$

where the points $a$, $b$, $c$ and $d$ are as shown in Figure 2.2.

– *Gaussian*

$$\mu_A(x) = \exp\left(-\frac{(x-c)^2}{2\sigma^2}\right)$$

where $c$ is the centre and $\sigma$ is the width of the gaussian distribution.

- **Database** Contains linguistic variables and their associated membership functions. Each linguistic variable $X$ is characterised by

$$X = (x, T(x), U, G, M)$$

where $x$ is the name of the variable, $T(x)$ is the term set (e.g., {Low, Medium, High}), $U$ is the universe of discourse, $G$ is the syntactic rule, and $M$ is the semantic rule assigning meaning to each term [71].

- **Rule Base Contains the set of IF-THEN rules for the specific modelling problem**

**Rule Representation**

$$R_i : \text{IF } x_1 \text{ is } A_1 \text{ AND } x_2 \text{ is } A_2 \ldots \text{ AND } x_n \text{ is } A_n \text{ THEN } y \text{ is } B \tag{2.1}$$

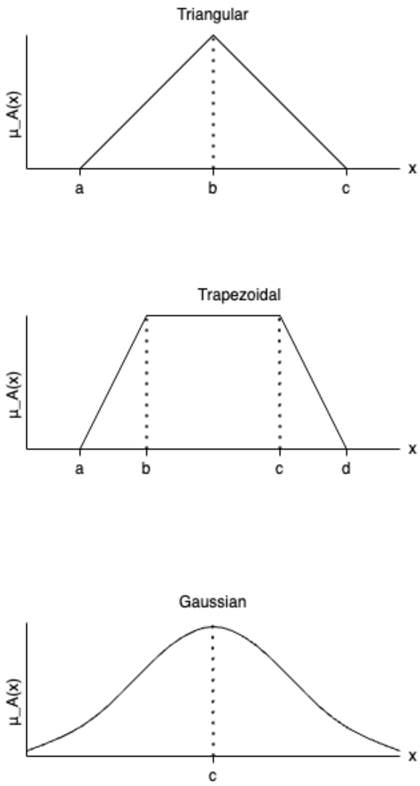


Figure 2.2: Commmon Type-1 Membership Functions

where:

- $R_i$ represents the $i$-th rule
- $x_1, x_2, \ldots, x_n$ are input linguistic variables
- $A_1, A_2, \ldots, A_n$ are antecedent fuzzy sets
- $y$ is the output linguistic variable
- $B$ is the consequent fuzzy set

For example, a sample rule is

$$\text{IF } (Temperature \text{ is } High) \text{ AND } (Humidity \text{ is } Low) \text{ THEN } (FanSpeed \text{ is } Fast)$$

where $Temperature$, $Humidity$ are input linguistic variables and $FanSpeed$ is the output linguistic variable. $High$, *Low* and *Fast* are fuzzy sets linked to their respective linguistic variables. Structurally, the rule consists of an antecedent (the IF part) and a consequent (the THEN part). During rule execution, an inference engine evaluates the antecedent's firing strength using a fuzzy operator (e.g., AND, typically implemented as the minimum or product of membership degrees). This strength is then applied to the consequent through an implication method to derive a qualified fuzzy output, which is later aggregated and defuzzified.

- **Fuzzy Inference Engine** Infers fuzzy output from inputs according to the rules in the rule base. The inference process involves:
  1. *Rule activation* Evaluate each rule's antecedent to determine its firing strength
  2. *Implication* Apply implication operator to obtain rule output
  3. *Aggregation* Combine outputs from all fired rules
- **Defuzzification Module** Maps fuzzy values to real-valued output domain using various techniques described below.

The database and rule base collectively form the Knowledge Base (KB).

**Fuzzy Inference Process - Illustrated Example** Consider a simple temperature control system with two inputs (temperature and humidity) and one output (fan speed).

**Step 1: Deriving Membership Functions from Data** Historical observations (e.g., temperature, humidity, fan speed) are analysed to define linguistic categories. For each variable, data-driven methods (such as clustering or histogram-based partitioning) are used to obtain membership functions:

- Apply a partitioning method (e.g., fuzzy c-means) to identify typical value ranges
- Fit corresponding membership functions (e.g., triangular or trapezoidal) to each cluster
- Assign labels such as *Cold, Warm, Hot* for temperature and *Low, Medium, High* for humidity

**Step 2: Fuzzification** Given crisp inputs Temperature = 25°C, Humidity = 60%

$$\mu_{Cold}(25) = 0, \quad \mu_{Warm}(25) = 0.7, \quad \mu_{Hot}(25) = 0.3$$
$$\mu_{Low}(60) = 0.2, \quad \mu_{Medium}(60) = 0.8, \quad \mu_{High}(60) = 0$$

**Step 3: Rule Evaluation** Example rules:

- $R_1$: IF Temperature is Warm AND Humidity is Medium THEN Fan Speed is Medium
- $R_2$: IF Temperature is Hot AND Humidity is Low THEN Fan Speed is High

$$\alpha_1 = \min(0.7, 0.8) = 0.7, \qquad \alpha_2 = \min(0.3, 0.2) = 0.2$$

**Step 4: Implication**

$$\mu^{R_1}_{Medium}(y) = \min(0.7, \mu_{Medium}(y))$$
$$\mu^{R_2}_{High}(y) = \min(0.2, \mu_{High}(y))$$

**Step 5: Aggregation**

$$\mu_{output}(y) = \max(\mu^{R_1}_{Medium}(y),\ \mu^{R_2}_{High}(y))$$

**Common Defuzzification Techniques** The defuzzification process converts the aggregated fuzzy output into a crisp value. Common techniques include

1. **Centre of Gravity (COG)/ Centroid Method**

$$y^* = \frac{\int_Y y \cdot \mu_{output}(y)\, dy}{\int_Y \mu_{output}(y)\, dy} \tag{2.2}$$

   For discrete universe:

$$y^* = \frac{\sum_{i=1}^{n} y_i \cdot \mu_{output}(y_i)}{\sum_{i=1}^{n} \mu_{output}(y_i)} \tag{2.3}$$

2. **Centre of Area (COA) / Bisector Method** Finds the vertical line that divides the area under the membership function into two equal parts

$$\int_{y_{min}}^{y^*} \mu_{output}(y)\, dy = \int_{y^*}^{y_{max}} \mu_{output}(y)\, dy$$

3. **Mean of Maximum (MOM)**

$$y^* = \frac{1}{|M|} \sum_{y \in M} y$$

   where $M = \{y | \mu_{output}(y) = \max_z \mu_{output}(z)\}$

4. **Centre of Sums Method**

$$y^* = \frac{\sum_{i=1}^{n} \int y \cdot \mu_i(y)\, dy}{\sum_{i=1}^{n} \int \mu_i(y)\, dy}$$

   where $\mu_i(y)$ is the membership function of the consequent of the $i$th rule.

**Comparison of Defuzzification Methods** Each defuzzification method has distinct characteristics [72]:

- **COG** is widely used and produces smooth outputs, but it requires more computation.
- **MOM** is fast and easy to apply, although it can lead to output discontinuities.
- **COA** uses area information to determine the output, but this increases the computational colpmexity.
- **Weighted Average** is efficient in Sugeno systems, provided that the consequents are singletons.

**Complete Fuzzy Inference Example** For the temperature control example, using COG defuzzification

$$y^* = \frac{\int_0^{100} y \cdot \max(\min(0.7, \mu_{Medium}(y)), \min(0.2, \mu_{High}(y)))\, dy}{\int_0^{100} \max(\min(0.7, \mu_{Medium}(y)), \min(0.2, \mu_{High}(y)))\, dy}$$

If the output membership functions are triangular with: - Medium: centre at 50, support in the range [30, 70] - High: centre at 80, support in the range [60, 100].

The defuzzified output would be approximately 58% fan speed.

### 2.2.3 Interval Type-2 Fuzzy Rule-Based Systems

Interval Type-2 (IT2) fuzzy sets extend type-1 fuzzy sets by introducing a footprint of uncertainty (FOU), which models uncertainties such as linguistic ambiguity, noisy data, or imprecise measurements more effectively [69]. In IT2-FRBSs, each membership grade is represented as an interval rather than a single value, and the secondary membership function specifies the degree of belief associated with every point in that interval. In the interval type-2 case, this secondary membership function is typically uniform and usually is assigned the value 1 across the entire interval, indicating that all values within the interval are considered equally plausible representations of the underlying uncertainty.

Structurally, IT2-FRBSs are similar to type-1 FRBSs - with a few key differences. The fuzzifier in an IT2-FRBS sends type-2 input fuzzy sets to the inference engine and the inference engine outputs type-2 output fuzzy sets, which need to be type-reduced to type-1 fuzzy sets before the defuzzification into a crisp output. The structure is as shown in Figure 2.3. IT2-FRBSs are particularly suited for real-world applications because they effectively handle higher levels of uncertainty and imprecision inherent in complex environments, providing superior performance over traditional type-1 systems [73]. Their robustness to noise stems from the use of IT2 fuzzy sets that model uncertainties in data and expert knowledge more accurately, while the flexibility in rule design allows for adaptable linguistic rules that can be easily tuned to accommodate varying real-world scenarios without requiring precise mathematical models [73].

#### Mathematical Formulation

An IT2 fuzzy set $\tilde{F}$ is defined as:

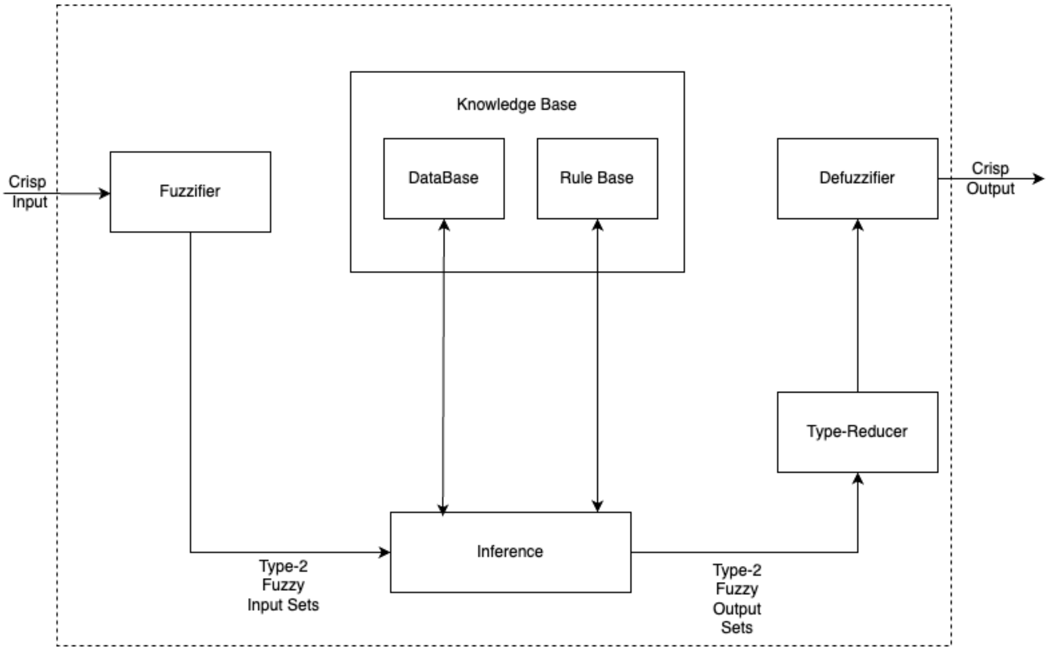


Figure 2.3: Structure of a Type-2 Fuzzy Rule Based System

$$\tilde{F} = \int_{x \in X} \left[ \int_{u \in [\underline{\mu}_{\tilde{F}}(x), \overline{\mu}_{\tilde{F}}(x)]} 1/u \right] /x \tag{2.4}$$

where $\underline{\mu}_{\tilde{F}}(x)$ and $\overline{\mu}_{\tilde{F}}(x)$ are the lower and upper membership functions respectively, defining the FOU. The FOU represents the uncertainty in the primary membership function and is bounded by the following as shown in Figure 2.4:

- Upper Membership Function (UMF): $\overline{\mu}_{\tilde{F}}(x)$

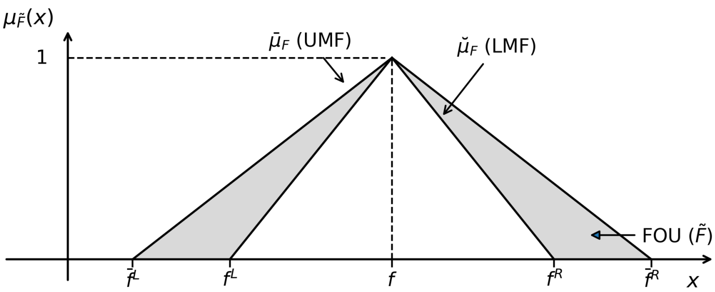


Figure 2.4: Interval Type-2 Fuzzy Set

- Lower Membership Function (LMF): $\underline{\mu}_{\tilde{F}}(x)$

For computational efficiency, the secondary membership function is assumed to be uniform (equal to 1) within the interval $[\underline{\mu}_{\tilde{F}}(x), \overline{\mu}_{\tilde{F}}(x)]$, which characterises the IT2 nature.

**Operations on IT2 Fuzzy Sets**

The following operations define how IT2 fuzzy sets are combined during rule evaluation and inference in IT2-FRBSs, using standard Zadeh operators (maximum for union, minimum for intersection) [72]. Key operations for IT2 fuzzy sets include:

- **Union**:

$$\underline{\mu}_{\tilde{F}\cup\tilde{G}}(x) = \underline{\mu}_{\tilde{F}}(x) \vee \underline{\mu}_{\tilde{G}}(x) \tag{2.5}$$

$$\overline{\mu}_{\tilde{F}\cup\tilde{G}}(x) = \overline{\mu}_{\tilde{F}}(x) \vee \overline{\mu}_{\tilde{G}}(x) \tag{2.6}$$

- **Intersection**:

$$\underline{\mu}_{\tilde{F}\cap\tilde{G}}(x) = \underline{\mu}_{\tilde{F}}(x) \wedge \underline{\mu}_{\tilde{G}}(x) \tag{2.7}$$

$$\overline{\mu}_{\tilde{F}\cap\tilde{G}}(x) = \overline{\mu}_{\tilde{F}}(x) \wedge \overline{\mu}_{\tilde{G}}(x) \tag{2.8}$$

- **Complement**:

$$\underline{\mu}_{\neg\tilde{F}}(x) = 1 - \overline{\mu}_{\tilde{F}}(x) \tag{2.9}$$

$$\overline{\mu}_{\neg\tilde{F}}(x) = 1 - \underline{\mu}_{\tilde{F}}(x) \tag{2.10}$$

$\vee$ and $\wedge$ represent maximum and minimum operators respectively. A typical rule in an IT2-FRBS is formulated as:

$$R_i: \ IF \ x_1 \ is \ \tilde{A}_{i1} \ AND \ \ldots \ AND \ x_n \ is \ \tilde{A}_{in} \ THEN \ y \ is \ \tilde{B}_i$$

where $\tilde{A}_{ij}$ are IT2 fuzzy sets for the antecedents corresponding to input variables $x_j$, and $\tilde{B}_i$ is the IT2 fuzzy set for the consequent (in Mamdani-type systems). For Takagi-Sugeno-Kang (TSK)-type IT2-FRBSs, the consequent is typically a linear function with interval parameters:

$$R_i: \; IF\ x_1\ is\ \tilde{A}_{i1}\ AND\ \ldots\ AND\ x_n\ is\ \tilde{A}_{in}\ THEN\ y = [y_i^l, y_i^r]$$

where $[y_i^l, y_i^r]$ represents the interval-valued output, often computed as $y_i^l = \sum_{j=0}^{n} c_{ij}^l x_j$ and $y_i^r = \sum_{j=0}^{n} c_{ij}^r x_j$ with $x_0 = 1$.

**Common Type-Reduction Techniques**

In type-2 FRBS, the aggregated output is a type-2 fuzzy set, which is first reduced to a type-1 fuzzy set (or interval) before defuzzification. Common type-reduction techniques include [72, 74]:

1. **Centroid Type-Reduction - Karnik-Mendel (KM) Algorithm**: Computes an interval $[y_l, y_r]$ where:

$$y_l = \frac{\sum_{i=1}^{L} y_i \bar{\mu}_{output}(y_i) + \sum_{i=L+1}^{N} y_i \underline{\mu}_{output}(y_i)}{\sum_{i=1}^{L} \bar{\mu}_{output}(y_i) + \sum_{i=L+1}^{N} \underline{\mu}_{output}(y_i)} \quad (2.11)$$

   (and similarly for $y_r$ with switched upper/lower), with the switch point $L$ found iteratively. In 2.11, the switch point $L$ determines the index at which the algorithm changes from using upper membership values $\bar{\mu}_{output}$ to lower membership values $\underline{\mu}_{output}$. It is derived iteratively: starting from an initial estimate, the output values $y_i$ are ordered, the centroid is recomputed, and $L$ is updated until the index stabilises, indicating convergence to the correct switch between upper and lower membership contributions.

2. **Enhanced Karnik-Mendel (EKM)**: An optimised version of KM with faster convergence [75].

3. **Nie-Tan Method**: A simplified approximation [76]:

$$\mu_{reduced}(y) = \underline{\mu}_{output}(y) + 0.5(\bar{\mu}_{output}(y) - \underline{\mu}_{output}(y))$$

4. **Wu-Mendel Uncertainty Bounds**: Provides bounds on the centroid without full iteration [77].

After type-reduction, defuzzification (e.g., average of $[y_l, y_r]$) yields the crisp output.

**Type-2 Fuzzy Inference Process - Example**

Continuing with the example presented for type-1 FRBS, consider a simple temperature control system with two inputs (temperature and humidity) and one output (fan speed), now using an IT2-FRBS to handle uncertainty in the membership functions:

- **Fuzzification:** Given crisp inputs: Temperature = 25°C, Humidity = 60For temperature (interval memberships):

$$[\underline{\mu}Cold(25), \bar{\mu}Cold(25)] = [0, 0]$$
$$[\underline{\mu}Warm(25), \bar{\mu}Warm(25)] = [0.6, 0.8]$$
$$[\underline{\mu}Hot(25), \bar{\mu}Hot(25)] = [0.2, 0.4]$$

  For humidity:

$$[\underline{\mu}Low(60), \bar{\mu}Low(60)] = [0.1, 0.3]$$
$$[\underline{\mu}Medium(60), \bar{\mu}Medium(60)] = [0.7, 0.9]$$
$$[\underline{\mu}High(60), \bar{\mu}High(60)] = [0, 0]$$

- **Rule Evaluation:** Example rules (same as type-1, but with type-2 sets):
  - $R_1$: IF Temperature is Warm AND Humidity is Medium THEN Fan Speed is Medium
  - $R_2$: IF Temperature is Hot AND Humidity is Low THEN Fan Speed is High

  Firing intervals using minimum operator (AND):

$$\begin{aligned}[\underline{\alpha}_1, \bar{\alpha}_1] &= [\min(\underline{\mu}_{\text{Warm}}(25), \underline{\mu}_{\text{Medium}}(60)), \min(\bar{\mu}_{\text{Warm}}(25), \bar{\mu}_{\text{Medium}}(60))] \\ &= [\min(0.6, 0.7), \min(0.8, 0.9)] \\ &= [0.6, 0.8] \\ [\underline{\alpha}_2, \bar{\alpha}_2] &= [\min(\underline{\mu}_{\text{Hot}}(25), \underline{\mu}_{\text{Low}}(60)), \min(\bar{\mu}_{\text{Hot}}(25), \bar{\mu}_{\text{Low}}(60))] \\ &= [\min(0.2, 0.1), \min(0.4, 0.3)] \\ &= [0.1, 0.3]\end{aligned}$$

- **Step 3: Implication** Using minimum implication on intervals (applied to upper and lower membership functions of consequents):

$$\begin{aligned}
\bar{\mu}^{R_1}_{\text{Medium}}(y) &= \min(\bar{\alpha}_1, \bar{\mu}_{\text{Medium}}(y)) \\
&= \min(0.8, \bar{\mu}_{\text{Medium}}(y)) \\
\underline{\mu}^{R_1}_{\text{Medium}}(y) &= \min(\underline{\alpha}_1, \underline{\mu}_{\text{Medium}}(y)) \\
&= \min(0.6, \underline{\mu}_{\text{Medium}}(y)) \\
\bar{\mu}^{R_2}_{\text{High}}(y) &= \min(\bar{\alpha}_2, \bar{\mu}_{\text{High}}(y)) \\
&= \min(0.3, \bar{\mu}_{\text{High}}(y)) \\
\underline{\mu}^{R_2}_{\text{High}}(y) &= \min(\underline{\alpha}_2, \underline{\mu}_{\text{High}}(y)) \\
&= \min(0.1, \underline{\mu}_{\text{High}}(y))
\end{aligned}$$

- **Step 4: Aggregation** Using maximum operator on intervals:

$$\begin{aligned}
\bar{\mu}_{\text{output}}(y) &= \max(\bar{\mu}^{R_1}_{\text{Medium}}(y), \bar{\mu}^{R_2}_{\text{High}}(y)) \\
\underline{\mu}_{\text{output}}(y) &= \max(\underline{\mu}^{R_1}_{\text{Medium}}(y), \underline{\mu}^{R_2}_{\text{High}}(y))
\end{aligned}$$

**Complete Type-2 Fuzzy Inference Example** For the temperature control example, assume the output membership functions are IT2 triangular with:

Medium: UMF support range in [25, 75] with apex at 50, LMF support range in [35, 65] with apex at 50. High: UMF support range in [55, 100] with apex at 80, LMF support range in [65, 95] with apex at 80.

Using KM type-reduction as shown in equation 2.11 followed by defuzzification (average of centroids): The defuzzified output is approximately 57% fan speed.

### 2.2.4 Takagi-Sugeno-Kang Fuzzy Systems

Fuzzy logic systems provide a framework for modelling complex, non-linear systems by incorporating human-like reasoning through linguistic rules. Unlike traditional Mamdani fuzzy systems, which utilise fuzzy sets in both antecedents and consequents, TSK fuzzy systems employ crisp functional consequents, offering computational efficiency and ease of optimisation. This section discusses type-1 and type-2 TSK fuzzy systems, highlighting their structures, inference mechanisms, and applications in handling uncertainty.

**Type-1 TSK Fuzzy Systems**

Type-1 TSK fuzzy systems, introduced by Takagi and Sugeno in 1985, represent a hybrid approach combining fuzzy logic with linear regression [61]. The system comprises a set of rules where the antecedents are fuzzy propositions, and the consequents are crisp linear functions of the inputs.

In a type-1 TSK fuzzy system with $M$ rules and inputs $\mathbf{x} = [x_1, x_2, \ldots, x_n]^T$. The $i$-th rule is expressed as:

$$R^i: \ IF\ x_1\ is\ A_1^i\ AND\ \ldots\ AND\ x_n\ is\ A_n^i\ THEN\ y^i = p_0^i + p_1^i x_1 + \cdots + p_n^i x_n,$$

where $A_j^i$ denotes the type-1 fuzzy set for the $j$-th input in the $i$-th rule, and $\mathbf{p}^i = [p_0^i, p_1^i, \ldots, p_n^i]^T$ are the consequent parameters.

The firing strength of the $i$-th rule, $\mu^i(\mathbf{x})$, is computed using a T-norm (usually minimum or product). Using product as T-norm:

$$\mu^i(\mathbf{x}) = \prod_{j=1}^{n} \mu_{A_j^i}(x_j),$$

where $\mu_{A_j^i}(x_j)$ is the membership degree of $x_j$ in $A_j^i$.

The overall output $y$ is obtained via a weighted average defuzzification:

$$y = \frac{\sum_{i=1}^{M} \mu^i(\mathbf{x}) y^i}{\sum_{i=1}^{M} \mu^i(\mathbf{x})}.$$

This formulation allows type-1 TSK systems to approximate any continuous function on a compact domain [78]. They are useful in control and modelling applications due to their interpretability in the antecedent part and ability to integrate expert knowledge with data-driven parameter tuning via methods such as least squares.

However, type-1 fuzzy sets assume precise membership functions, which may not adequately capture uncertainties arising from noisy data, linguistic ambiguities, or varying expert opinions [72].

**Type-2 TSK Fuzzy Systems**

To address the limitations of type-1 systems, type-2 TSK fuzzy systems extend the framework by incorporating type-2 fuzzy sets, which possess fuzzy membership functions themselves. This enables modelling of higher-order uncertainties.

In an IT2 TSK system, the antecedents are IT2 fuzzy sets, denoted as $\tilde{A}^i_j$, characterised by upper and lower membership functions, $\bar{\mu}_{\tilde{A}^i_j}(x_j)$ and $\underline{\mu}_{\tilde{A}^i_j}(x_j)$. The consequents remain linear functions, similar to type-1.

The firing interval for the $i$-th rule is:

$$[\underline{\mu}^i(\mathbf{x}), \bar{\mu}^i(\mathbf{x})] = \left[\prod_{j=1}^{n} \underline{\mu}_{\tilde{A}^i_j}(x_j), \prod_{j=1}^{n} \bar{\mu}_{\tilde{A}^i_j}(x_j)\right].$$

Inference involves type-reduction to obtain a type-1 fuzzy set, followed by defuzzification. A common method is the KM algorithm, as discussed in Section 2.2.3, which computes the centroid endpoints. The crisp output is then $y = (y_l + y_r)/2$, where $y_l$ and $y_r$ are computed using 2.11.

Type-2 TSK systems excel in environments with significant uncertainty, such as noisy sensor data or imprecise measurements [79]. Nonetheless, they incur higher computational costs due to type-reduction.

Although TSK systems provide computational efficiency and enhanced approximation capabilities, they are generally less explainable and interpretable than Mamdani systems due to their use of crisp linear functions in rule consequents, which lack the linguistic transparency of Mamdani's fuzzy sets [80]. Bringing the two model types together in a single structure is one of the key focus areas of this research.

### 2.2.5 Genetic and Evolutionary Fuzzy Systems

Genetic Fuzzy Systems (GFSs) represent a significant advancement in FRBSs, employing evolutionary algorithms such as genetic algorithms [81] and genetic programming [82] for learning purposes. These systems address the challenge of designing FRBSs as a search problem, utilising the ability of evolutionary algorithms to explore large search spaces for near-optimal solutions [69].

**Architecture and Learning Approaches**

GFS learning is divided into two main categories:

- **Learning Knowledge Base (KB) components**: This includes generating and selecting fuzzy rules, defining database components (e.g., membership function shapes and fuzzy set numbers), and co-learning rules and database components simultaneously.

- **Parameter tuning**: This involves optimising membership function parameters, inference system parameters, and defuzzification weights to improve system performance.

The first systematic tuning approach was proposed by Thrift [83], while rule selection was pioneered by Ishibuchi et al. [84]. Learning database components was first addressed by Cordón et al. [62]. These methods balance solution quality with computational cost, often requiring careful design to avoid overfitting.

**Evolutionary Approaches for IT2-FRBSs**

Evolutionary algorithms have proven particularly effective for optimising IT2-FRBSs due to their ability to handle complex search spaces and multiple objectives simultaneously. These approaches address various aspects of system design, from parameter tuning to comprehensive rule base construction:

- **Parameter Learning**

  Evolutionary algorithms excel at learning IT2-FRBS parameters, including membership function parameters and FOU boundaries:

  - **Genetic algorithm-based parameter tuning**: Santoso et al. [85] developed methods to learn parameters and FOU from scratch. The chromosome representation for an IT2 fuzzy set includes:

    $$\text{Chromosome} = [c^L, c^U, \sigma^L, \sigma^U, w_1, w_2, \ldots, w_n]$$

    where $c^L, c^U$ are the lower and upper centres, $\sigma^L, \sigma^U$ are the spreads, and $w_i$ are rule weights.

– **Differential evolution**: Employs mutation and crossover operations specifically designed for continuous parameter optimisation:

$$v_i = x_{r1} + F \cdot (x_{r2} - x_{r3})$$

where $F$ is the scaling factor and $x_{r1}, x_{r2}, x_{r3}$ are randomly selected individuals.

– **Particle swarm optimisation**: Updates particle velocities based on:

$$v_i(t+1) = w \cdot v_i(t) + c_1 \cdot r_1 \cdot (p_{best} - x_i(t)) + c_2 \cdot r_2 \cdot (g_{best} - x_i(t)),$$

where $v_i(t+1)$ is the velocity of particle $i$ at time $t+1$, $v_i(t)$ is the velocity at time $t$, $w$ is the inertia weight, $c_1$ is the cognitive coefficient, $r_1$ is a random number between 0 and 1, $p_{best}$ is the personal best position of particle $i$, $x_i(t)$ is the position of particle $i$ at time $t$, $c_2$ is the social coefficient, $r_2$ is a random number between 0 and 1, and $g_{best}$ is the global best position. Positions are then updated as $x_i(t+1) = x_i(t) + v_i(t+1)$.

- **Structure Optimisation**

The structure of IT2-FRBSs, including the number of rules and their organisation, significantly impacts system performance:

– **Multi-objective ant colony optimisation**: Su et al. [86] employed Ant Colony Optimisation (ACO) to simultaneously optimise accuracy and interpretability. The pheromone update rule is:

$$\tau_{ij}(t+1) = (1-\rho) \cdot \tau_{ij}(t) + \sum_{k=1}^{m} \Delta\tau_{ij}^{k}$$

where $\rho$ is the evaporation rate and $\Delta\tau_{ij}^{k}$ is the pheromone deposited by ant $k$.

– **Hierarchical structure evolution**: Evolutionary algorithms determine optimal hierarchical arrangements to reduce rule explosion in high-dimensional problems.

– **Variable selection**: Genetic algorithms identify relevant input variables, reducing system complexity whilst maintaining performance.

- **Rule Base Design and Selection**

  Rule selection represents a critical aspect of IT2-FRBS design, as it directly impacts both system performance and interpretability. Evolutionary approaches offer sophisticated mechanisms for rule generation and selection:

  **Genetic Programming for Rule Generation**

  – **Tree-based representation**: Rules are encoded as expression trees, allowing flexible rule structures:

$$\text{Fitness}(R) = w_1 \cdot \text{Accuracy}(R) + w_2 \cdot \text{Interpretability}(R) - w_3 \cdot \text{Complexity}(R)$$

  – **Grammar-guided genetic programming**: Ensures syntactically correct IT2 fuzzy rules through formal grammar constraints.

  **Rule Selection Strategies**

  – **Michigan approach**: Each individual in the population represents a single rule. The final rule base is formed by selecting the best individuals:

$$RB_{final} = \{R_i | \text{Fitness}(R_i) > \theta \text{ and } i \in \{1, 2, \ldots, N_{pop}\}\}$$

  – **Pittsburgh approach**: Each individual represents an entire rule base. This allows for global optimisation but requires larger computational resources:

$$\text{Fitness}(RB) = \frac{\text{Accuracy}(RB)}{\text{Size}(RB)^{\alpha}} \cdot \text{Coverage}(RB)$$

  where $\alpha$ controls the penalty for large rule bases.

  – **Iterative rule learning (IRL)**: Rules are learned sequentially, with each new rule covering instances not adequately covered by existing rules:

$$\text{Coverage}(R_i) = \frac{|\{x \in D | R_i \text{ fires for } x\}|}{|D|}$$

**Advances in GFS**

Recent developments in Genetic Fuzzy Systems (GFS) extend the field across several research areas:

1. **Hybrid and autonomous learning frameworks** Hybrid GFS frameworks integrate genetic optimisation with complementary learning paradigms, including neural networks, co-evolution, and autonomous adaptation mechanisms. Neuro-genetic fuzzy models combine neural parameter learning with evolutionary optimisation of fuzzy structures [87], while co-evolutionary designs have been explored for interpretable reinforcement learning policies [88]. More recently, autonomous learning fuzzy systems emphasise minimal human intervention by enabling online structure evolution, rule generation, and parameter adaptation driven directly by data streams [89]. Particle-swarm-optimised and self-adaptive fuzzy systems further demonstrate how evolutionary optimisation can be embedded within continuous learning processes rather than confined to offline training [90, 91].

2. **Enhanced uncertainty modelling** To improve robustness under noise, ambiguity, and non-stationarity, GFS research has expanded beyond type-1 fuzzy representations. Interval type-2 fuzzy systems explicitly model uncertainty through upper and lower membership bounds [92], and genetic optimisation has been employed to tune both primary and secondary membership parameters for improved resilience in noisy environments [86]. Complementarily, belief-driven and evidential fuzzy inference systems introduce alternative uncertainty handling mechanisms that support adaptive rule confidence estimation and improved generalisation in evolving classification tasks [93].

3. **Structural and representation evolution** Modern GFS increasingly evolves *model structure* in addition to tuning parameters, including the number of rules, membership function shapes, and linguistic partitions. Grammar-based and genetic programming (GP) approaches enable the automatic discovery of compact and interpretable fuzzy

rule bases [94]. Recent multilayer and hierarchical fuzzy architectures further extend this idea by allowing the evolution of deep fuzzy representations, where multiple fuzzy layers cooperate to capture complex, high-dimensional patterns while preserving interpretability at the rule level [95, 96]. Dimensionality compression mechanisms embedded within evolving fuzzy neural networks demonstrate how structural evolution can mitigate the curse of dimensionality without sacrificing transparency.

4. **Interpretability and explainability** Aligning GFS with explainable AI (XAI) has become a central research direction, with interpretability treated as an explicit optimisation objective rather than a secondary by-product. Interpretability metrics, linguistic constraints, and rule semantics are increasingly incorporated into evolutionary fitness functions [80]. Evolving and multilayer fuzzy systems further contribute to this goal by maintaining human-understandable rule bases while supporting adaptive learning, thereby bridging the gap between *transparent reasoning* and *scalable performance* in complex data-driven applications [89, 95].

### 2.2.6 Challenges, Limitations and Future Directions

Despite significant advances in FRBSs, several fundamental challenges and limitations persist, particularly in classical Mamdani implementations:

**Structural and Scalability Challenges**

1. **Interpretability-Accuracy Trade-off**: Whilst FRBSs are interpretable by design, increasing complexity for improved accuracy often compromises interpretability [69]. This fundamental tension remains a critical consideration in system design.

2. **Scalability and Curse of Dimensionality**: The number of possible rules grows exponentially with increasing variables:

$$|R_{max}| = \prod_{i=1}^{n} |T_i|$$

where $|T_i|$ is the number of linguistic terms for the $i$th variable. For 10 variables with 5 terms each, this yields $5^{10} \approx 9.77$ million potential rules, presenting significant challenges for big data applications.

3. **Lack of Flexibility in Partition Design**: Fixed membership functions cannot adapt to changing data distributions, with predetermined partition granularity limiting adaptability. For instance, in control systems, fixed membership functions may fail to capture shifts in sensor data distributions over time [97].

**Performance and Accuracy Limitations**

4. **Inconsistent Rule Outputs in Logical Gaps**: In certain subsets of the input space, the rule base may fail to produce outputs that align with expected directional relationships. These gaps occur when adjacent input vectors trigger outputs that move in the wrong direction or differ by implausible margins due to incomplete rule coverage or coarse partitioning, undermining system reliability and user trust [98, 99].

5. **Limited Accuracy in Regression Tasks**: The reliance on linguistic rules and defuzzification in Mamdani FRBSs can lead to reduced accuracy compared to functional models like Takagi-Sugeno-Kang systems. Fixed membership functions and coarse partitioning may fail to capture complex non-linear relationships, whilst defuzzification methods (e.g., centroid) introduce approximation errors, particularly in noisy or high-dimensional datasets [99].

The inconsistencies in rule outputs and the limited accuracy of Mamdani FRBSs represent two of the key focal points of this research, motivating the development of more adaptive and accurate approaches whilst maintaining the interpretability advantages inherent to fuzzy systems.

### 2.2.7 Conclusion

FRBSs, particularly Mamdani-type systems and their IT2 extensions, continue to evolve through integration with genetic and evolutionary algorithms. These hybrid approaches offer promising solutions for tackling complex classification and regression problems, while preserving the interpretability crucial for transparent decision-making domains. Improving predictive accuracy without sacrificing semantic consistency, interpretability, and explainability remains an active research area - this work contributes directly to addressing these challenges.

## 2.3 Fuzzy Time Series

Fuzzy Time Series (FTS) represents a significant advancement in forecasting methodologies, particularly for handling linguistic and uncertain data. This approach was pioneered by Song and Chissom in 1993, who introduced a novel framework for analysing time series data where observations are represented as fuzzy sets rather than crisp numerical values [32, 100].

Traditional time series analysis methods require precise numerical data and often assume specific statistical properties. However, many real-world scenarios involve linguistic descriptions or imprecise observations. There have been many methods to forecast continuous values in the literature. However, none of them could be applied when the historical data are linguistic values. Song and Chissom's FTS methodology addressed this limitation by incorporating fuzzy set theory into time series analysis.

### 2.3.1 Basic Definitions

Song and Chissom [100] established the foundational definitions for FTS:

**Universe of Discourse** Let $U = \{u_1, u_2, ..., u_n\}$ be the universe of discourse, which is a finite set containing all possible values of the time series.

**Fuzzy Time Series** Let $Y(t)$ $(t = ..., 0, 1, 2, ...)$ be a subset of $\mathbb{R}$, the universe of discourse on which fuzzy sets $f_i(t)$ are defined. If $F(t)$ is a collection of $f_i(t)$ $(i = 1, 2, ...)$, then $F(t)$ is called a FTS defined on $Y(t)$.

**Fuzzy Logical Relationship** Suppose $F(t)$ is implied by $F(t-1)$ only, that is, $F(t) = F(t-1) \circ R(t, t-1)$. Then this relation can be expressed as $F(t-1) \rightarrow F(t)$, where $R(t, t-1)$ is the fuzzy relationship between $F(t)$ and $F(t-1)$, and is called the first order model of $F(t)$. If $F(t)$ is caused by $F(t-1)$ only, denoted as $F(t-1) \rightarrow F(t)$, then this relationship $R(t, t-1)$ is called a first-order fuzzy logical relationship.

**Time-Invariant vs Time-Variant** If for any $t$, $R(t, t-1)$ is independent of $t$, that is, for any $t$, $R(t, t-1) = R$, then $F(t)$ is called a time-invariant FTS, otherwise it is called a time-variant FTS.

### 2.3.2 Mathematical Framework

The core of Song and Chissom's methodology involves fuzzy relational equations. For a first-order time-invariant model, the relationship is expressed as:

$$F(t) = F(t-1) \circ R,$$

where $\circ$ denotes the max-min composition operation, and $R$ is the fuzzy relational matrix. The max-min composition is defined as:

$$\mu_{F(t)}(u_j) = \max_{u_i \in U}[\mu_{F(t-1)}(u_i) \wedge \mu_R(u_i, u_j)],$$

where $\mu_{F(t)}(u_j)$ represents the membership degree of $u_j$ in $F(t)$.

### 2.3.3 Time-Invariant Model

Song and Chissom [100] developed a systematic approach for forecasting using FTS:

**Step 1: Define the Universe of Discourse**

- Determine the range $[D_{min} - D_1, D_{max} + D_2]$ where $D_{min}$ and $D_{max}$ are the minimum and maximum values in the historical data.
- $D_1$ and $D_2$ are positive numbers chosen to ensure proper coverage.

**Step 2: Partition the Universe of Discourse**

- Divide $U$ into equal-length intervals $u_1, u_2, ..., u_n$.
- Define linguistic variables $A_1, A_2, ..., A_n$ corresponding to each interval.

**Step 3: Define Fuzzy Sets** For each linguistic variable $A_i$, define its membership function. Song and Chissom used triangular membership functions:

$$\mu_{A_i}(x) = \begin{cases} \frac{x-a_{i-1}}{a_i-a_{i-1}} & \text{if } a_{i-1} \leq x \leq a_i \\ \frac{a_{i+1}-x}{a_{i+1}-a_i} & \text{if } a_i < x \leq a_{i+1} \\ 0 & \text{otherwise} \end{cases}$$

**Step 4: Fuzzify Historical Data** Convert each historical observation into a fuzzy set based on maximum membership principle.

**Step 5: Establish Fuzzy Logical Relationships** Identify all relationships of the form $A_i \to A_j$ from the fuzzified data.

**Step 6: Create the Fuzzy Relational Matrix** Construct the fuzzy relational matrix R as the union (max) of all individual fuzzy relations:

$$R = \max_k (F(t_k - 1) \times F(t_k)),$$

where $\times$ is the min-based Cartesian product. In discretized form,

$$R_{pq} = \max_t \min(\mu_{F(t-1)}(u_p), \mu_{F(t)}(u_q))$$

.

**Step 7: Forecast** Apply the max-min composition to obtain fuzzy forecasts, then defuzzify to get crisp values.

### 2.3.4 Time-Variant Model

Song and Chissom [101] extended their methodology to develop time-variant models for forecasting. The key distinction between time-invariant and time-variant models lies in the treatment of the fuzzy relationship matrix. Whilst time-invariant models assume a constant relationship matrix $R$, time-variant models allow the fuzzy relationship matrix to change over time:

$$F(t) = F(t-1) \circ R(t)$$

This formulation captures the evolving nature of relationships in dynamic systems. Song and Chissom [101] employed neural networks for defuzzification in their time-variant model.

### 2.3.5 Fuzzy Logical Relationship Groups

A major limitation of Song and Chissom's approach was computational complexity, since their forecasting process relied on repeated max–min composition operations across fuzzy matrices. Chen introduced a more efficient method [102] by replacing those operations with simplified arithmetic calculations and a clearer representation of fuzzy logical relationships (FLRs).

Chen's notation takes the form

$$A_i \to A_{j1}, A_{j2}, \ldots, A_{jk},$$

where $A_i$ denotes the antecedent fuzzy set (the current state) and $A_{j1}, A_{j2}, \ldots, A_{jk}$ denote the possible consequent fuzzy sets (the next states). Instead of generating a separate relationship for each transition, Chen groups all relationships that share the same antecedent:

$$\text{If} \quad A_i \to A_{j1}, A_i \to A_{j2}, \ldots, A_i \to A_{jk}$$

$$\text{then group as} \quad A_i \to (A_{j1}, A_{j2}, \ldots, A_{jk}).$$

This notation is significant for two reasons. First, it reduces the number of operations required during forecasting, replacing matrix-based max–min compositions with direct arithmetic on the grouped consequents. Second, it introduces a structured way to represent multiple future states from a single past state, which makes the model easier to interpret and scale to larger datasets. As a result, Chen's method improved both computational efficiency and practical applicability, and became the foundation for most subsequent fuzzy time-series forecasting models.

### 2.3.6 Simplified Forecasting Rules

Chen's forecasting rules:

1. If $F(t-1) = A_i$ and $A_i \to A_j$, then $F(t) = A_j$
2. If $F(t-1) = A_i$ and $A_i \to A_{j1}, A_{j2}, ..., A_{jk}$, then:

   $$F(t) = \frac{1}{k}(A_{j1} + A_{j2} + ... + A_{jk})$$

   where the sum is computed on the midpoints of the intervals associated with each $A_{jk}$, yielding a crisp forecast.
3. If $F(t-1) = A_i$ and there is no relationship starting from $A_i$, then $F(t) = A_i$

This approach eliminated the need for complex matrix operations whilst maintaining forecasting accuracy.

### 2.3.7 Subsequent Developments

#### High-Order Models

Chen (2002) studied high-order FTS models of forecasting to overcome the drawbacks of first order FTS models and implemented it to forecast the student enrollments of the University of Alabama. High-order models consider multiple lagged variables:

$$F(t) = f(F(t-1), F(t-2), ..., F(t-n))$$

where $n$ is the order of the model.

#### Interval Partitioning Improvements

Several researchers addressed the critical issue of universe of discourse partitioning:

- **Huarng's Ratio-Based Approach (2001)** Huarng [103] proposed using distribution-based intervals rather than equal-length partitions, improving forecast accuracy for non-uniformly distributed data.

- **Entropy-Based Partitioning** Building on information theory, Tsaur, Yang, and Wang [104] employ an entropy-based measure to quantify the degree of fuzziness inherent in fuzzy relations used within fuzzy time series models. In their approach, entropy is computed from the membership values of fuzzy sets or fuzzy relation matrices and is used as an evaluation criterion for selecting an appropriate time-invariant fuzzy relation matrix. Lower entropy indicates a more informative and less ambiguous fuzzy relation structure. By incorporating entropy into fuzzy relation analysis, the method provides a principled means of assessing and comparing fuzzy relations, thereby improving the reliability of fuzzy time series forecasting.

#### Weighted Fuzzy Time Series

Yu [105] introduced weights to fuzzy logical relationships based on their recurrence:

$$w_{ij} = \frac{c_{ij}}{\sum_k c_{ik}}$$

where $c_{ij}$ represents the count of transitions from $A_i$ to $A_j$.

#### Heuristic Models

Huarng [106] developed heuristic models by integrating problem-specific heuristic knowledge with Chen's model to improve forecasting. These models incorporated trend information into forecasting rules, considering whether the time series is increasing, decreasing, or stable.

### 2.3.8 Further Advances

#### Machine Learning Integration

Later developments have integrated machine learning techniques with FTS methods:

- Neural networks for fuzzy relationship identification [107]
- Genetic algorithms for optimal interval partitioning [30]
- Support vector machines for defuzzification [108]
- Particle swarm optimisation for parameter selection [109]
- There has been a move towards hybrid fuzzy models with other approaches such as LSTMs. Targeted at improving forecasting accuracy, these approaches are inherently more complex and reduce interpretability [110].
- Use of non-stationary fuzzy sets to update membership parameters [111, 112].

#### Type-2 Fuzzy Time Series

Extensions to type-2 fuzzy sets provide additional capability for handling uncertainty in membership functions themselves. This approach allows for better representation of uncertainty when the membership degrees themselves are uncertain [113].

#### Multivariate Fuzzy Time Series

Lee et al. [114] extended FTS to multivariate cases. In a multivariate fuzzy inference process, an m-factor k-order FTS is a time series consisting of m factors that are of kth order. This allows for modelling complex relationships between multiple variables, significantly expanding the applicability of FTS methods to real-world problems involving multiple interrelated time series.

### 2.3.9 Limitations of Fuzzy Time Series Approaches

Despite its widespread adoption and success, the FTS methodology exhibits some fundamental limitations. The use of the maximum membership principle for defuzzification fundamentally contradicts the core tenet of fuzzy set theory. As Zadeh [1] originally conceived, fuzzy sets allow elements to belong to multiple sets simultaneously with varying degrees of membership, capturing the inherent ambiguity in real-world phenomena. However, the maximum membership principle employed in most FTS models forces a crisp classification by selecting only the fuzzy set with highest membership degree, thereby discarding valuable information about partial memberships [102]. This reductionist approach potentially undermines the very advantage that fuzzy logic promises over traditional methods.

Another critical limitation concerns the selection of optimal look-back periods or time steps in model construction. Whilst Song and Chissom [32] proposed first-order models and subsequent researchers extended to high-order models [115], the field lacks a systematic methodology for determining the optimal order or look-back window. This is particularly problematic given that time series data typically exhibits strong autocorrelation at nearby time steps, making the choice of model order crucial for capturing relevant temporal dependencies without introducing noise from irrelevant historical data. The absence of principled approaches for order selection often leads to ad hoc choices based on trial and error, potentially compromising model performance. Furthermore, empirical evidence suggests that forecasting accuracy in FTS models often improves with an increased number of fuzzy sets used for partitioning the universe of discourse [103]. However, this performance gain comes at a significant cost to model interpretability and explainability. As the number of fuzzy sets increases, the linguistic interpretation of each set becomes less meaningful, and the model begins to approximate a complex non-linear function rather than providing intuitive linguistic rules. This raises the question of whether a Mamdani-style FRBS might offer a more principled approach, maintaining the balance between accuracy and interpretability through explicit if-then rules that preserve the linguistic nature of the fuzzy sets [29]. Such systems could potentially address both the maximum membership limitation and the interpretability issue by utilising the full fuzzy output and providing transparent reasoning processes.

An added challenge in FTS approaches is the subjectivity inherent in partitioning the universe of discourse and defining fuzzy rules, which introduces bias and hampers reproducibility. The fuzzification process, including the determination of the number of fuzzy sets and interval lengths, often relies on expert judgment or arbitrary choices, leading to variability in model outcomes across different implementations [116].

FTS models also face significant challenges when dealing with volatile data, where statistical properties like mean and variance change over time. Real-world series, such as financial markets or environmental data, often exhibit trends, seasonality, or heteroskedasticity, leading to poor performance without additional preprocessing like differencing or decomposition. To address this, extensions like non-stationary FTS with time-varying parameters or fuzzy-wavelet methods have been proposed to adapt fuzzy sets dynamically based on residuals or distribution changes [111]. However, these adaptations often require hybrid integrations, highlighting the inherent limitation of base FTS in handling dynamic environments.

Additionally, computational demands and error accumulation pose substantial limitations, particularly in high-order models or long-term forecasts. As model complexity increases with more fuzzy sets or higher orders, time and memory requirements grow quadratically, making FTS inefficient for large datasets. [117]. This is exacerbated in multivariate or high-dimensional scenarios, where rule explosion further amplifies computational burdens.

Hybrid approaches integrating FTS with machine learning techniques, such as neural networks, LSTM, or random forests, have mitigated some of these limitations by automating partitioning, handling data distribution drift through adaptive learning, and reducing computational overhead via optimized architectures. For example, hybrids using ML for fuzzy relation modeling or interval optimization improve accuracy in volatile data while addressing subjectivity [118,119]. However, these integrations introduce new issues, such as overfitting, especially in high-order multivariate hybrids where complex fuzzy rules overfit training data, leading to poor generalization on unseen test periods [120]. Techniques like regularization or ensemble methods can alleviate overfitting, but they add further complexity to model design and interpretation.

This research contributes by providing improvements to time-step selection targeted specifically towards fuzzy models, as well as enhancements to the FTS interpretability and explainability.

## 2.4 Classification and Regression Metrics

Performance evaluation is fundamental to machine learning, providing quantitative measures to assess model quality and compare different algorithms. The choice of appropriate metrics depends on the task type, data characteristics, and specific business requirements. This section provides an overview of the most commonly used performance metrics for classification and regression tasks as relevant to the work presented in subsequent chapters, including their mathematical formulations and practical interpretations. This overview is not an exhaustive overview of classification and regression metrics, for example, multi-class metrics have not been presented as the research presented does not contain any examples of multi-class classification tasks.

### 2.4.1 Classification Metrics

Classification tasks involve predicting discrete class labels. The performance metrics for classification are typically derived from the confusion matrix, which tabulates the correct and incorrect predictions for each class.

#### Confusion Matrix

For binary classification, the confusion matrix is a 2×2 table

| | | **Predicted** | |
|---|---|---|---|
| | | Positive | Negative |
| **Actual** | Positive | TP | FN |
| | Negative | FP | TN |

Table 2.1: Binary Classification Confusion Matrix

where,

- TP (True Positives): Correctly predicted positive cases
- TN (True Negatives): Correctly predicted negative cases
- FP (False Positives): Incorrectly predicted positive cases (Type I error)
- FN (False Negatives): Incorrectly predicted negative cases (Type II error)

**Accuracy**

Accuracy measures the fraction of predictions that are correct

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

While intuitive, accuracy can be misleading for imbalanced datasets where one class significantly outnumbers others.

**Precision**

Precision (also called Positive Predictive Value) measures the fraction of positive predictions that are actually correct

$$\text{Precision} = \frac{TP}{TP + FP}$$

High precision indicates low false positive rate, meaning when the model predicts positive, it is usually correct.

**Recall (Sensitivity)**

Recall (also called Sensitivity or True Positive Rate) measures the fraction of actual positive cases that are correctly identified

$$\text{Recall} = \frac{TP}{TP + FN}$$

High recall indicates low false negative rate, meaning the model captures most of the actual positive cases.

**Specificity**

Specificity (True Negative Rate) measures the fraction of actual negative cases that are correctly identified

$$\text{Specificity} = \frac{TN}{TN + FP}$$

**ROC Curve, AUC, and Gini Coefficient**

The Receiver Operating Characteristic (ROC) curve plots the True Positive Rate (Recall) against the False Positive Rate at various threshold settings

$$\text{False Positive Rate} = \frac{FP}{FP+TN} = 1 - \text{Specificity}$$

The Area Under the ROC Curve (AUC-ROC) provides a single scalar value representing the model's ability to discriminate between classes. AUC ranges from 0 to 1, where,

- AUC = 1.0: Perfect classifier
- AUC = 0.5: Random classifier
- AUC = 0.0: Perfectly wrong classifier

The Gini coefficient, derived from the ROC curve, is another metric frequently used to evaluate model performance. It is calculated directly from the AUC as

$$\text{Gini} = 2 \times \text{AUC}_{\text{ROC}} - 1$$

The Gini coefficient scales the AUC from [0.5, 1] to a range of [0, 1], making it easier to interpret model performance relative to random chance. Specifically

- Gini = 1.0: Perfect classifier
- Gini = 0.0: Random classifier

Thus, the Gini coefficient complements the AUC by providing an intuitive measure of model effectiveness.

### 2.4.2 Regression Metrics

Regression tasks involve predicting continuous numerical values. Let $y_i$ represent the true values and $\hat{y}_i$ represent the predicted values for $i = 1, 2, \ldots, n$ samples.

**Mean Absolute Error (MAE)**

MAE measures the average absolute difference between predicted and actual values

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i|$$

MAE is robust to outliers and provides an intuitive interpretation in the same units as the target variable.

**Mean Squared Error (MSE)**

MSE measures the average squared difference between predicted and actual values

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

MSE penalizes larger errors more heavily due to the squaring operation, making it sensitive to outliers.

**Root Mean Squared Error (RMSE)**

RMSE is the square root of MSE, providing error measurement in the same units as the target variable

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2} = \sqrt{\text{MSE}}$$

**R-squared (Coefficient of Determination)**

R-squared measures the proportion of variance in the dependent variable explained by the model. A residual $e_i$ is the difference between the observed value $y_i$ and the predicted value $\hat{y}_i$:

$$e_i = y_i - \hat{y}i$$

These residuals represent the error for each individual data point. When we square these differences and sum them up, we get the Sum of Squares of Residuals ($SSres$), which is a key component in calculating $R^2$.

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} = 1 - \frac{SS_{res}}{SS_{tot}},$$

where $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ is the mean of actual values, $SS_{res}$ is the sum of squares of residuals, and $SS_{tot}$ is the total sum of squares.

$R^2$ ranges from 0 to 1 for meaningful models, where 1 indicates perfect fit and 0 indicates the model performs no better than predicting the mean.

**Adjusted R-squared**

Adjusted R-squared accounts for the number of predictors in the model

$$R_{adj}^2 = 1 - \frac{(1-R^2)(n-1)}{n-p-1},$$

where $p$ is the number of predictors. This metric penalizes the addition of irrelevant predictors.

## 2.5 Summary

This chapter provides foundational concepts essential for understanding the thesis work. Section 2.2 introduces Fuzzy Rule-Based Systems (FRBSs), covering their architecture including fuzzification, knowledge base structure, inference mechanisms, and defuzzification methods, along with detailed explanations of Mamdani-type systems and various membership function types. The chapter then reviews fuzzy time series prediction approaches in Section 2.3, including classical methods like Song and Chissom's work, Chen's simplified approach, and modern hybrid techniques combining fuzzy systems with neural networks and genetic algorithms, while also discussing interval-valued and type-2 fuzzy approaches that handle higher-order uncertainty. Section 2.4 concludes with comprehensive coverage of classification and regression evaluation metrics which are used throughout the thesis to assess experimental results.

# Chapter 3

# Enhancing Explainability for Fuzzy Time-Series Approaches

## 3.1 Introduction

This chapter proposes the Optimized Time Series Interval Type-2 Fuzzy System, a framework that addresses the limitations of existing fuzzy time-series methods discussed in Section 1.1.1 while maintaining their inherent explainability advantages. This approach makes several key contributions:

- **Automated Time Period Discovery** We introduce a correlation-based algorithm that automatically identifies the most predictive time-steps from the data, eliminating the need for manual specification and ensuring optimal temporal feature selection.

- **Optimized Fuzzy Time Sets** Our method employs a systematic optimization procedure to determine the optimal shape and parameters of fuzzy sets around selected time-steps, maximizing classification performance while maintaining interpretability.

- **Linguistically Meaningful Features** By aggregating multiple time-steps through fuzzy membership functions, we create features that align with human temporal reasoning (e.g., $recent\_price$ instead of $price(t-3)$), improving model interpretability. The approach presented in our research brings the explainability much closer to an intuitive understanding of temporal periods as discussed in [121].

- **Noise-Robust Representations** The fuzzy aggregation process naturally smooths out noise in individual observations, resulting in more stable and generalisable models compared to traditional approaches such as raw feature-based classifiers or deterministic rule-based systems that are highly sensitive to outliers and measurement errors.

- **Smaller Feature Space** Picking out the most predictive time-steps also means that the feature space is much smaller compared to other techniques which mine for complex features using transformations [122].

Our experimental results demonstrate that OTS-IVFS achieves competitive accuracy with state-of-the-art black-box methods while providing fully interpretable models. On the Earthquake dataset, our approach achieves a 16.67% improvement in accuracy over the best existing methods, while maintaining comparable performance on other benchmark datasets. More importantly, the generated models consist of human-readable rules that domain experts can easily validate and trust.

The remainder of this chapter is organised as follows: Section 3.2 provides a review of the related work and identifies gaps in current research. Section 3.3 presents our proposed OTS-IVFS methodology in detail. Section 3.4 describes our experimental setup and results. Finally, Section 3.6 concludes with a discussion of the limitations of our work and identifies further directions for research.

## 3.2 Literature Review

### 3.2.1 Evolution of Time Series Classification

Time series classification has evolved significantly over the past there decades. Historically, early approaches focused on nearest neighbour methods, distance measures, and similarity-based techniques due to their simplicity and effectiveness in handling temporal alignment and irregularity. Dynamic Time Warping (DTW) emerged as particularly influential for its robustness to variations in timing or speed within series [123, 124]. The field advanced rapidly following the establishment of the UCR Time Series Archive, providing standardised benchmarks for algorithm comparison [125].

Bagnall et al. [126] conducted the first comprehensive evaluation of time-series classification algorithms, categorising approaches into whole series, interval-based, shapelet-based, dictionary-based, and combination methods. This taxonomy structured the expanding field and highlighted the dominance of ensemble approaches. More recently, Middlehurst et al. (2024) provided an updated "bake off redux" reviewing and evaluating post-2017 algorithms, including new deep learning integrations and scalable methods, while confirming the continued superiority of ensembles on benchmark datasets.

Multivariate time-series classification introduces additional complexities compared to univariate analysis. Ruiz et al. [127] extended the original "bake off" to multivariate datasets, demonstrating that simple concatenation of univariate methods often fails to capture inter-dimensional dependencies.

### Traditional Machine Learning Extensions

Several researchers have extended traditional time-series learning methods to multivariate contexts. Schäfer and Leser [128] adapted the bag-of-patterns paradigm to capture cross-dimensional dependencies through the WEASEL+MUSE framework, extracting symbolic features from multiple dimensions and their derivatives. Although these approaches are generally more interpretable than deep learning models, the symbolic discretisation process obscures the relationship between original signal values and model predictions, limiting their practical transparency for domain experts. Subsequent refinements, such as enhanced symbolic aggregation approximation, primarily address discretisation artifacts and computational efficiency rather than advancing inherent explainability.

Shapelet-based methods have also been generalised to multivariate time series. Baydogan and Runger [129] proposed learning discriminative shapelets jointly across dimensions; however, the resulting patterns often lack clear semantic meaning despite their theoretical interpretability. Later work on interpretable shapelet learning, including adversarial regularisation strategies [130], improves robustness and stability but remains largely performance-oriented, offering limited insight into the underlying temporal phenomena.

Distance-based techniques, including multivariate extensions of Dynamic Time Warping (DTW), model inter-dimensional dependencies explicitly [130]. While effective for classification, such metric-based approaches provide minimal explanatory value, as similarity scores do not reveal which temporal structures drive predictions [131]. Recent variants incorporate attention mechanisms to highlight influential alignments and reduce computational complexity, yet these additions fall short of delivering semantically grounded explanations.

Feature-based approaches extract statistical descriptors such as mean, variance, entropy, and autocorrelation across dimensions. Although straightforward, these global statistics often fail to capture localised or position-specific temporal patterns, such as transient events or recurring motifs within constrained time windows [132]. Moreover, expanding feature sets through extended lookback periods leads to feature explosion, further complicating interpretation. Semantics-preserving dimensionality reduction methods, such as the rough and fuzzy-rough set approaches proposed by Jensen and Shen [133], attempt to address this issue by reducing dimensionality while maintaining decision-relevant information. However, their application to complex multivariate time-series remains limited, and interpretability gains are often indirect.

More recently, feature grouping and selection strategies have been proposed to improve both scalability and transparency. Zheng et al. [134] introduced a graph-based feature grouping framework that identifies correlated and semantically related feature subsets prior to selection. While this approach offers improved structural insight into large feature spaces, its reliance on engineered features means that explanations remain one step removed from the original time-series values.

Embedding-based methods for multivariate time-series classification alleviate feature explosion by learning compact latent representations. Despite their effectiveness in boosting classification accuracy, these embeddings typically lack semantic alignment with the original signals, limiting their explanatory utility.

Finally, dictionary-based systems convert multivariate series into sparse feature vectors through windowing and discretisation, followed by classification on word-frequency representations. The accumulation of transformation steps significantly reduces interpretability, making it difficult to trace model decisions back to original observations [135, 136]. Adaptive windowing and scalable variants improve efficiency but do not substantially address the fundamental explainability limitations inherent in these pipelines.

### 3.2.2 Explainable Time Series Classification

The demand for explainable AI has driven research into interpretable time-series classification methods, following post-hoc explanations of opaque models and inherently interpretable approaches.

#### Post-hoc Explanation Methods

Assaf and Schumann [137] used gradient-based attribution to explain CNN predictions on multivariate series. While offering visual insights, explanations remain disconnected from domain knowledge and lack actionability. Updated techniques incorporating saliency maps with temporal constraints improve visualization stability, but focus on reliability without advancing domain-aligned interpretability.

Siddiqui et al. [138] developed TSViz for visualizing neural attention in time-series classification, highlighting important time-steps but failing to explain significance or inter-series interactions. Recent multimodal extensions, such as InstrucTime for instruction-tuned attention, mitigate interaction oversights through natural language queries, though these enhancements prioritize flexible reasoning over core explainability.

LIME (Local Interpretable Model-agnostic Explanations) based adaptations for time-series analysis [139] perturb inputs for local understanding, but defining meaningful temporal perturbations is challenging, with explanations lacking stability [140]. Stabilization via ensemble perturbations addresses variability, but mainly enhances robustness without deeper interpretability.

#### Inherently Interpretable Methods

Decision trees and variants offer natural interpretability but struggle with temporal dependencies [126]. Time Series Forests [141] use statistical features and temporal importance curves, though large node counts diminish inspectability [142]. Randomization-based extensions improve scalability, but without boosting explainability beyond base structures.

Rule-based methods, including fuzzy and non-fuzzy variants, form a key interpretable category. Fuzzy approaches to time-series prediction are discussed in Section 2.3. For classification, recent non-fuzzy work extracts temporal rules from learned feature spaces. Dempster et al. [143] introduced ROCKET (RandOm Convolutional KErnel Transform) for high-dimensional convolutional features, with Detach-ROCKET (Detached RandOm Convolutional KErnel Transform) [144] reducing features up to 98% via Sequential Feature Detachment for concise explanations. However, rules remain grounded in abstract kernels, constraining intuitive access for experts. Extensions like COCALITE (COmmon CAuse LInear TErm) [145], combining Catch22 (Canonical Time-series CHaracteristics (22 selected features)) [146] features with lightweight networks, mitigate computational intensity, but primarily for performance without explainability advances.

Shapelet-based methods classify via similarity to discriminative subsequences [147], achieving comparable performance to opaque models while preserving transparency [148]. Yet, shapelets often lack semantic correspondence with domain phenomena, creating interpretation gaps despite visualization potential [149]. Support Vector Machine ensemble extensions enhance discriminative power, but focus on accuracy over semantic alignment.

#### Fuzzy Approaches to Time Series Classification

Fuzzy logic handles temporal uncertainty and enables linguistic interpretation. Song and Chissom's [32] foundational fuzzy time-series targeted forecasting, but extensions apply to classification. Traditional methods like Chen's simplified relationships [102] and weighted variants [105] map observations to single sets, limiting uncertainty handling.

Type-2 fuzzy systems better manage uncertainty via interval memberships [73, 150]. For time-series classification, recent works like fuzzy-probabilistic representation learning combine probabilistic weights with classifiers for robust features, though applications emphasize accuracy over interpretability. Škrjanc et al. [151] introduced evolving type-2 systems for streaming data, focusing on adaptation but with limited structural interpretability. A 2024 method combining time-series features and fuzzy memberships enhances global behaviour capture in classification, mitigating noise sensitivity via membership integration, however, it prioritizes generalization without advancing linguistic explainability [152]. Fuzzy attention-integrated transformers generate learnable fuzziness scores for temporal importance, addressing attention opacity, but extensions improve forecasting performance rather than classification explainability [153].

### 3.2.3 Recent Advances and Hybrid Approaches

Current research combines approaches for balanced performance and interpretability.

#### Ensemble Methods

The 2024 "bake off redux" confirms the superior performance of ensemble models, with superior performance from HIVE-COTE [154] as well as HIVE-COTE 2.0 [155], which incorporated diverse features. However, the ensemble's complexity comes at the cost of interpretation. Another high performing approach, MultiROCKET [156] generates convolutional features, again, with abstract nature limiting interpretability.

#### Neural-Symbolic Integration

Recent work has explored extracting temporal logic rules from neural networks through differentiable frameworks [157, 158], offering a bridge between symbolic reasoning and deep learning. However, the derived rules often exhibit high complexity, limiting their practical interpretability. Alternative approaches, such as neuro-fuzzy systems [159], demonstrate promising trade-offs between transparency and predictive performance by generating human-readable rules, though their application to time-series classification remains under-explored.

Meanwhile, the adaptation of large language models to time-series data has introduced novel architectures, particularly through transformer-based tokenization schemes that enable multimodal reasoning [160]. The models have reported high accuracy, but the architecture of these models does not support explainability and interpretability.

### 3.2.4 Gaps in Current Literature

#### Research Gaps

Our analysis reveals several fundamental limitations in current approaches:

- **Temporal Abstraction** Existing methods predominantly operate on either individual time-steps or fixed window segments, lacking the capacity for flexible, hierarchical temporal abstractions that better align with human cognitive processes in pattern recognition.

- **Multivariate Interpretability** While accuracy benchmarks continue to improve, there remains insufficient attention to developing interpretable representations of cross-variable relationships, particularly in non-fuzzy rule-based systems where transparency is crucial.

- **Linguistic Temporal Patterns** Current approaches largely fail to generate truly intuitive linguistic descriptions of temporal patterns. Even fuzzy methods typically anchor their explanations to numerical time-step references rather than semantically meaningful temporal periods (e.g., ”early morning” vs. timestamps).

- **Domain Adaptation** Configuring interpretable systems currently requires substantial domain expertise. The development of adaptive learning frameworks that can automatically discover meaningful temporal abstractions while preserving interpretability remains an open challenge.

OTS-IVFS addresses these via optimised linguistic abstractions, leveraging fuzzy advantages for enhanced multivariate and temporal interpretability. Future work must balance accuracy, interpretability, and efficiency for domain-useful insights.

## 3.3 Proposed Method: Optimized Time Series Interval Type-2 Fuzzy System

### 3.3.1 Overview

We propose the Optimized Time Series Interval Type-2 Fuzzy System (OTS-IVFS), which transforms time series data into linguistically interpretable features through optimized fuzzy time periods. Unlike traditional approaches that treat each time-step as an independent feature, our method aggregates temporal information using fuzzy sets defined over meaningful time intervals, resulting in more robust and explainable models.

### 3.3.2 Time-Based Fuzzification of Time Series

Traditional time series classification approaches create features directly from individual time-steps, resulting in feature vectors like $(x(t-1), x(t-2), \ldots, x(t-n))$. This leads to high-dimensional feature spaces that are prone to noise and difficult to interpret. We propose a fundamentally different approach that groups time-steps into fuzzy linguistic periods.

#### Creating Linguistic Time Periods

Consider a time series $X = (x_1, x_2, \ldots, x_T)$ where each prediction uses $n$ preceding observations as historical context. Instead of using individual time-steps, we define fuzzy sets over the temporal dimension to represent linguistic concepts such as “Recent”, “Medium Term”, and “Long Term”.

For example, the membership function for "Recent" might be defined as:

$$\mu_{Recent}(t) = \begin{cases} 1 & \text{if } 1 \leq t \leq 3 \\ \frac{7-t}{7-3} & \text{if } 3 < t \leq 7 \\ 0 & \text{otherwise} \end{cases} \tag{3.1}$$

where $t$ represents the temporal lag from the current time point.

**Feature Extraction through Fuzzy Aggregation**

Given a fuzzy time period $F$ with membership function $\mu_F(t)$, we create a time-based feature by aggregating the time-series values weighted by their membership degrees:

$$x_F = \frac{\sum_{t=1}^{n} x(t) \cdot \mu_F(t)}{\sum_{t=1}^{n} \mu_F(t)} \tag{3.2}$$

This transformation converts a sequence of raw values into a single interpretable feature. For instance, if we have two input time series (price_a and price_b) and three fuzzy time periods, we generate six features:

- price_a_Recent
- price_a_Medium_Term
- price_a_Long_Term
- price_b_Recent
- price_b_Medium_Term
- price_b_Long_Term

### 3.3.3 Optimization of Time Periods

A key contribution of this approach is the automatic selection of time-steps and fuzzy set parameters using a data-driven approach.

**Automatic Time Step Selection**

The automatic time-step selection algorithm identifies the most predictive time-steps by analysing the correlation structure between the time series and the target variable. This approach differs fundamentally from traditional time-series analysis techniques, as it searches for well-separated time-steps that maximise predictive power whilst avoiding redundancy. The algorithm operates through a systematic process.

First, it calculates the correlation between each time-step of the input time series and the target variable across the entire dataset. This produces a correlation curve that reveals the temporal relationship between historical values and the prediction target. Following the correlation calculation, the algorithm fits polynomials of increasing degree to the correlation time series. Starting with low-degree polynomials, the algorithm progressively increases the polynomial order until the fitted curve adequately captures the correlation pattern or if some pre-defined maximum degree of polynomial is reached. The fitting process terminates when polynomial predictions fall outside the observed correlation range. The algorithm then identifies local maxima in the fitted polynomial curve between adjacent minima. These maxima represent time-steps with the highest predictive correlation whilst maintaining separation from neighbouring peaks. For negative correlations, the algorithm inverts the correlation values and repeats the process.

This method offers several advantages over conventional approaches. Unlike the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) commonly used in econometric literature for lag selection, this method does not require the assumption of stationarity [161] and is robust to outliers, which otherwise distort the second-order moments used in classical correlation measures [162]. Additionally, it can handle classification problems where the output variable is not directly related to the time series values because the method does not depend on Autocorrelation. The algorithm automatically determines the optimal number of time-steps without requiring user specification for the number of time-steps to be selected as discussed in the Automated Time Step Selection Algorithm in section 3.3.5. For example, in the epilepsy dataset, applying this method to 178 time-steps yielded seven key time points: 0, 10, 60, 82, 94, 157, and 174. These selected time-steps form the foundation for subsequent fuzzy set construction, ensuring that the most informative temporal patterns are captured whilst maintaining computational efficiency.

1. **Correlation Analysis**: Calculate the correlation $\rho(t)$ between each lagged time-step and the target variable:

$$\rho(t) = \text{corr}(X_t, Y) \tag{3.3}$$

2. **Polynomial Fitting**: Fit polynomials of increasing degree $d$ (where $d$ is capped to a maximum user-defined value) to smooth the correlation curve:

$$\hat{\rho}(t) = \sum_{i=0}^{d} a_i t^i \tag{3.4}$$

3. **Time Step Identification**: Select time-steps corresponding to local maxima between consecutive minima in $\hat{\rho}(t)$. This ensures selected time-steps are well-separated and capture different temporal patterns.

### Fuzzy Set Parameter Optimization

For each selected time-step $t_s$, the system optimizes the fuzzy set parameters by testing various configurations within a search range. The optimization process identifies the optimal type-1 fuzzy set that maximizes the classification performance for the selected time-step by building a simple linear model on the model features related to the selected time-step in the training data.

$$\mathscr{F}_{t_s} = \arg\max_{F \in \mathscr{S}} \text{ROC-AUC}(F, \mathscr{D}_{train}) \tag{3.5}$$

where $\mathscr{S}$ represents the space of possible fuzzy sets centered around $t_s$ with varying widths, and $\mathscr{D}_{train}$ is the training dataset.

The search space $\mathscr{S}$ encompasses three primary fuzzy set types: Triangular, Trapezoidal, and Left/Right-Shoulder, as described in section 2.2.2.

The optimization algorithm employs a systematic search strategy. For a given maximum width parameter $w_{\max}$, the algorithm constructs all permissible fuzzy set configurations within the range $[t_s - w_{\max}, t_s + w_{\max}]$. Each configuration undergoes evaluation using the ROC-AUC metric on the training dataset.

The selection criterion prioritises fuzzy sets that capture meaningful temporal patterns whilst maintaining robustness to noise. The algorithm evaluates each candidate fuzzy set independently for each selected time-step. This approach ensures that the temporal characteristics specific to each time-step receive appropriate representation through optimally configured fuzzy sets.

The resulting optimized fuzzy sets transform the raw time series data into linguistically interpretable features. These features subsequently serve as inputs to the IT2 FRBS. The optimization process enhances both the interpretability and performance of the classification system by identifying fuzzy set configurations that effectively capture the underlying temporal patterns in the data.

### 3.3.4 Building the Interval Type-2 Fuzzy Rule-Based System

Once the optimized time-based features are extracted, we construct an IT2-FRBS for classification.

#### Fuzzification of Feature Values

Each time-based feature $x_F$ is fuzzified using IT2 fuzzy sets. The membership function of an IT2 fuzzy set $\tilde{A}$ is defined as:

$$\mu_{\tilde{A}}(x) = \int_{u \in [\underline{\mu}_{\tilde{A}}(x), \overline{\mu}_{\tilde{A}}(x)]} 1/u \tag{3.6}$$

where $\overline{\mu}_{\tilde{A}}$ and $\underline{\mu}_{\tilde{A}}$ represent the upper and lower membership functions.

#### Rule Generation and Optimization

The system generates IF-THEN rules of the form:

$$R_j : \text{IF } x_1 \text{ is } \tilde{A}_1^j \text{ AND } \ldots \text{ AND } x_F \text{ is } \tilde{A}_F^j \text{ THEN class is } C_j \tag{3.7}$$

Rule bases are optimized using genetic algorithms to maximise classification performance while maintaining interpretability by limiting the number of antecedents per rule.

#### Classification through Voting

The final classification is determined by aggregating votes from all fired rules:

$$V_{C_k}(x) = \sum_{R_j \in RB, C_j = k} \mu_{A_j}(x) \times RW_j \tag{3.8}$$

where $RW_j$ is the weight of rule $R_j$ based on its support and confidence.

### 3.3.5 Complete OTS-IVFS Algorithm

The complete OTS-IVFS algorithm integrates automated time-step selection, fuzzy set optimization, and IT2-FRBS construction into a unified framework. We present the algorithm in three phases for clarity.

---

**Algorithm 1** Phase 1: Automated Time Step Selection

**Require:** Time series dataset $\mathscr{D} = ((X^{(i)}, y^{(i)}))_{i=1}^{N}$, look-back window $n$
**Ensure:** Selected time-steps $T_{selected}$

**for** each input time series $k$ in dataset **do**
  Calculate correlation vector $\rho_k(t) = \text{corr}(X_k(t), Y)$ for $t = 1, \ldots, n$
  $\rho_{pos} \leftarrow$ positive values of $\rho_k(t)$
  $\rho_{neg} \leftarrow$ absolute values of negative correlations in $\rho_k(t)$
  **for** $\rho$ in $\{\rho_{pos}, \rho_{neg}\}$ **do**
    Fit polynomial of optimal degree $d^*$ to $\rho$
    Find all local minima $M = (m_1, m_2, \ldots, m_p)$ in fitted polynomial
    **for** each consecutive pair $(m_i, m_{i+1})$ in $M$ **do**
      $t_{max} \leftarrow \arg\max_{t \in [m_i, m_{i+1}]} \hat{\rho}(t)$
      Add $t_{max}$ to selected time-steps $T_k$
**return** $T_{selected} \leftarrow \bigcup_k T_k$

---

**Algorithm 2** Phase 2: Time-Based Fuzzy Set Optimization

**Require:** Selected time-steps $T_{selected}$, maximum width $w_{max}$, training data
**Ensure:** Optimized fuzzy sets $\{\mathscr{F}_{t_s}^*\}$

**for** each selected time-step $t_s \in T_{selected}$ **do**
  Generate candidate fuzzy sets with varying widths and shapes
    (Triangle, Trapezoid, Left/Right-Shoulder)
  **for** each candidate fuzzy set $F$ **do**
    Extract time-based features using $F$
    Calculate ROC-AUC score on $T_{selected}$ features in the training data
  $\mathscr{F}_{t_s}^* \leftarrow$ fuzzy set with highest ROC-AUC
**return** $\{\mathscr{F}_{t_s}^*\}$ for all $t_s \in T_{selected}$

---

**Algorithm 3** Phase 3: IT2-FRBS Construction and Optimization

**Require:** Optimized fuzzy sets $\{\mathscr{F}_{t_s}^*\}$, training/validation data
**Ensure:** Optimized IT2-FRBS classifier

Extract time-based features for all instances using $\{\mathscr{F}_{t_s}^*\}$
Create IT2 fuzzy sets for each feature using data distribution
Generate candidate rules (max 3 antecedents) with rule weights
Initialize genetic algorithm population of rule base configurations
**for** generation $g = 1$ to $G_{max}$ **do**
  Evaluate fitness on validation set
  Apply genetic operators (selection, crossover, mutation)
  Update best rule base if improved
**return** Optimized IT2-FRBS with linguistic rules

---

**Algorithm Complexity Analysis**

The computational complexity of OTS-IVFS can be analysed in three phases:

- **Time Step Selection**: Conditions: Correlations between each lagged input $X_k(t)$ and the target variable $Y$ are computed over $N$ training instances. Polynomial fitting is performed using least-squares regression with model selection over degrees 1 to $d_{\max}$. It is assumed that the look-back window size $n$ is significantly larger than the maximum polynomial degree ($n \gg d_{\max}$).

  Complexity: For each of the $K$ input time series, computing correlations across $n$ lags requires $O(N \cdot n)$ time. Polynomial fitting across candidate degrees incurs a cost of $O(n \cdot d_{\max}^3)$ as least-squares regression is performed for all degrees 1 to $d_{max}$ each requiring $O(n \cdot d^2)$ time. The cost of local extrema detection and peak selection is of lower order. Hence, the total complexity of Phase 1 is given by

  $$O\big(K \cdot (N \cdot n + n \cdot d_{\max}^3)\big)\,.$$

  Justification: In practice, $d_{\max}$ is small (typically $\leq 6$), and this phase produces a substantially reduced set of informative time-steps, limiting its overall computational impact.

- **Fuzzy Set Optimization**: Conditions: Fuzzy set optimisation is performed independently for each selected time-step. Each candidate fuzzy set is evaluated by computing membership values for all $N$ training instances. The number of fuzzy set shapes $S$ and the maximum width parameter $w_{\max}$ are treated as small constants. ROC-AUC computation is assumed to scale linearly with $N$ due to fixed-threshold or reused sorting strategies.

  Complexity: For each selected time-step $t_s \in T_{selected}$, the optimisation cost is

  $$O(S \cdot w_{\max} \cdot N).$$

Therefore, the total complexity of Phase 2 is

$$O(|T_{selected}| \cdot S \cdot w_{\max} \cdot N).$$

Justification: Although this phase scales linearly with the number of training instances, the number of selected time-steps satisfies $|T_{selected}| \ll n$ due to the automated selection in Phase 1.

- **Rule Base Optimization**: Conditions: Each fuzzy rule contains at most three antecedents. Rules are evaluated independently during inference. Interval type-2 fuzzy inference and type-reduction are implemented using standard centroid-based methods. Genetic algorithm parameters, including population size $|P|$ and maximum number of generations $G_{\max}$, are fixed.

  Complexity: Evaluating a single rule base configuration over $N$ instances requires $O(N \cdot |RB|)$ time. With $|P|$ candidate rule bases per generation and $G_{\max}$ generations, the total complexity of Phase 3 is

  $$O(G_{\max} \cdot |P| \cdot |RB| \cdot N).$$

  Justification: This phase dominates the overall computational cost. However, restricting the number of antecedents per rule and reducing the feature space through prior time-step selection significantly limits the effective rule base size.

While rule base optimisation is the most computationally demanding stage, the automated time-step selection reduces the feature dimensionality from $n \cdot K$ to $|T_{selected}| \cdot K$, where $|T_{selected}| \ll n$. This reduction offsets the additional costs introduced by fuzzy set optimisation and IT2-FRBS learning, making the OTS-IVFS framework computationally feasible for practical time-series classification tasks.

### 3.3.6 Example Application

Consider a binary financial time-series classification task in which the objective is to predict whether a stock's price trend on day $t+1$ will be *Upward* or *Downward*, based on historical price and volume data.

### Dataset Description

- **Input time series** ($K = 2$):
  1. Daily closing price
  2. Daily trading volume
- **Target variable**:
$$y^{(i)} = \begin{cases} 1, & \text{if } \mathrm{Close}_{t+1} > \mathrm{Close}_t, \\ 0, & \text{otherwise} \end{cases}$$
- **Look-back window**: $n = 60$ days
- **Number of training instances**: $N = 2000$

### Phase 1: Automated Time-Step Selection

For the input time series, correlations between lagged inputs and the target variable are computed as

$$\rho_k(t) = \mathrm{corr}(X_k(t), Y), \quad t = 1, \ldots, 60.$$

Polynomial fitting and local extrema detection identify the most informative temporal regions. For this dataset, the following time-steps are selected:

$$T_{\text{selected}} = \{3, 15, 45\}$$

These time-steps are shared across all input time series. Consequently, each input feature (e.g., price and volume) is evaluated at the same temporal locations. This reduces the original $60 \times 2 = 120$ lagged variables to 6 informative time-based features:

- *price_around_3*, *price_around_15*, *price_around_45*
- *volume_around_3*, *volume_around_15*, *volume_around_45*

### Phase 2: Time-Based Fuzzy Set Optimization

For each selected time-step, candidate fuzzy sets of different shapes (triangular, trapezoidal, left-shoulder, right-shoulder) and widths are evaluated using ROC-AUC on the training data.

Example: Time-Step $t = 3$

For the selected time-step $t = 3$, candidate fuzzy sets are generated over the normalized distributions of all input features (e.g., price and volume) observed at this lag. The same fuzzy set structure is applied to each feature evaluated at $t = 3$.

The candidate fuzzy sets considered include:

- **Triangular:** Triangle$[1,3,5]$
- **Triangular:** Triangle$[2,3,4]$
- **Trapezoidal:** Trapezoid$[1,2,4,5]$
- **Left-Shoulder:** Left-Shoulder$[0,0,2,4]$
- **Right-Shoulder:** Right-Shoulder$[3,5,7,7]$

Each candidate fuzzy set is evaluated by computing membership values for all training instances and all input features at time-step $t = 3$. The discriminative power of the resulting fuzzy features is assessed using the ROC-AUC metric, and the fuzzy set achieving the highest ROC-AUC is selected as the optimized fuzzy set for this time-step.

Using $T_{\text{selected}} = \{3, 15, 45\}$ and $K = 2$ input time series, the resulting fuzzy feature space consists of $|T_{\text{selected}}| \times K = 6$ time-based fuzzy features.

**Phase 3: IT2-FRBS Construction and Optimization**

Each time-based feature is partitioned into a fixed number of linguistic fuzzy sets (e.g., Low, Medium, and High). Interval type-2 fuzzy sets are then constructed for each linguistic term by introducing uncertainty in the membership functions to model data variability and noise.

Candidate fuzzy rules with up to three antecedents are generated and optimized using a genetic algorithm. Example rules include:

**R1**: IF price_around_15 is Low AND volume_around_3 is High
THEN trend is Upward

The optimized rule base obtained through genetic algorithm optimization consists of a small set of weighted, interpretable rules.

This end-to-end approach provides transparent reasoning while maintaining competitive classification accuracy, making it suitable for domains requiring both performance and explainability.

## 3.4 Experimental Evaluation

### 3.4.1 Experimental Design

This section presents the experimental evaluation of the Optimized Time Series Interval Valued Fuzzy System (OTS-IVFS). The experiments assess the system's performance across diverse time-series classification tasks whilst examining its interpretability advantages.

### 3.4.2 Datasets

Five datasets from the UCR and Time Series Classification repository were selected for evaluation [125, 163]. The datasets represent diverse application domains. Epileptic Seizures contains EEG readings for seizure detection. Earthquakes includes seismic data for earthquake prediction. SharePriceIncrease contains financial time-series data. ItalyPowerDemand represents electricity consumption patterns. SelfRegulationSCP1 contains brain-computer interface data [163]. Here's a brief description of the datasets:

1. **Epileptic Seizures:** EEG recordings from 500 individuals, initially comprising 4097 data points, are divided into 23 chunks of 178 data points each. The focus is binary classification to detect epileptic seizure cases versus non-seizure cases using labeled categories.

2. **Earthquakes:** Data is sourced from the Northern California Earthquake Data Center, consisting of hourly averaged readings from December 1st, 1967, to 2003. Major event prediction is framed as identifying readings over 5 on the Richter scale, not preceded by another major event for at least 512 hours, while negative cases involve readings below 4, preceded by at least 20 nonzero readings in the previous 512 hours. Segmentation, not sliding windows, is used. The dataset includes 368 negative cases and 93 positive cases among 86,066 hourly readings.

3. **Share Price Increase:** Derived from daily price movements of NASDAQ 100 companies, this dataset comprises 60-day data points representing percentage changes in

closing prices compared to the previous day. It focuses on predicting whether share prices will rise by more than 5% (Class 1) or not (Class 0) after quarterly EPS announcements. There are 1,931 cases: 1,326 in Class 0 and 605 in Class 1.

4. **Italy Power Demand:** Derived from twelve monthly electrical power demand time series from Italy, this dataset classifies days from October to March (inclusive) versus April to September.

5. **Self Regulation SCP1:** Involving healthy subjects controlling a cursor on a screen via slow cortical potentials, this dataset records EEG data from 6 channels at 256 Hz with a $\pm 1000\,\mu$V range. Training data includes 268 trials over 2 days, while test data (293 trials) has withheld labels.

### 3.4.3 Experimental Setup

#### Data Preparation

Each dataset includes predefined training and test splits. The training data was further divided into 70% for training, 15% for validation, and 15% for testing. The final results are reported on the original out-of-sample test sets (as provided by the data repository authors).

Feature engineering included calculating first-order differences and fixed differences from the primary time series. These derived features were denoted with prefixes 'Diff' and 'Fixed Diff' respectively.

#### Parameter Configuration

The OTS-IVFS framework was implemented using the parameter settings listed below. These values were selected to balance model expressiveness, interpretability, and computational feasibility.

- **Maximum fuzzy set width (5 time-steps):** The width of candidate fuzzy sets was limited to five time-steps to capture short- to mid-range temporal patterns while keeping the fuzzy set optimization process computationally manageable. Larger widths were found to increase the search space without providing significant performance gains.

- **Maximum rule length (3 antecedents):** The number of antecedents per fuzzy rule was restricted to three in order to preserve rule interpretability and avoid overly complex logical conditions. This constraint ensures that the resulting rule base remains comprehensible to human experts while still allowing meaningful feature interactions to be modeled.

- **Genetic algorithm population size (100):** A population size of 100 individuals was chosen to provide sufficient diversity in candidate rule bases, enabling effective exploration of the search space without incurring excessive computational cost.

- **Number of generations (50):** The number of generations was set large enough to allow convergence toward near-optimal rule bases. Empirically, increasing the number of generations beyond this value yielded diminishing performance improvements.

- **Crossover probability (0.8):** A crossover probability of 0.8 was adopted based on common practice in genetic algorithm implementations [164] and empirical experience with fuzzy rule base optimization. While higher crossover rates generally promote exploration of the solution space, this parameter could benefit from systematic hyperparameter tuning for optimal performance.

- **Mutation probability (0.1):** The mutation probability was set to 0.1 to maintain population diversity and prevent premature convergence. This value is higher than traditionally recommended for binary GAs [164], reflecting the increased need for exploration when optimizing complex symbolic structures such as fuzzy rule bases. Further hyperparameter optimization could refine this choice.

- **Number of fuzzy sets per feature (up to 6):** Allowing up to six fuzzy sets per feature provides sufficient linguistic granularity to model nonlinear relationships while avoiding excessive partitioning that could reduce interpretability and increase inference complexity.

#### Time Step Selection

The time-step selection algorithm identified predictive time periods through correlation analysis. For the Epilepsy dataset, the algorithm selected time-steps 0, 10, 60 , 82, 94, 157, and 174. figures 3.1 and 3.2 illustrates the correlation curve and polynomial fitting.

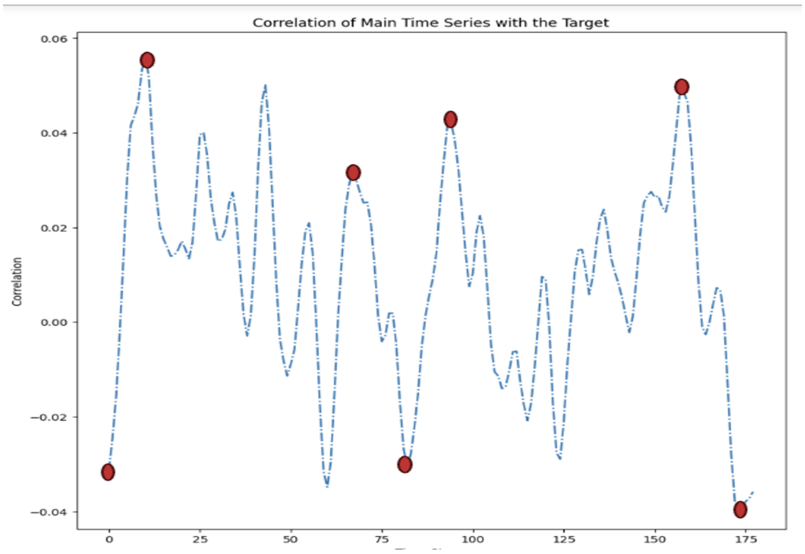


Figure 3.1: Correlation analysis for time-step selection on Epilepsy dataset

### Fuzzy Set Optimization

Time-based fuzzy sets were optimized around selected time-steps. The optimization process evaluated all permissible fuzzy set configurations within the maximum width parameter. ROC-AUC scores on the training set determined optimal configurations.

For the Epilepsy dataset, the optimization yielded:

- Time Step 0: Left-Shoulder[0,3,5]
- Time Step 10: Triangle[7,10,14]
- Time Step 60: Trapezoid[55,59,60,63]
- Time Step 82: Triangle[78,82,84]
- Time Step 94: Triangle[91,94,98]
- Time Step 157: Triangle[152,156,161]
- Time Step 174: Triangle[170,174,176]

## 3.4.4 Baseline Comparisons

OTS-IVFS performance was compared against the top performing algorithms for each dataset as reported in the "bake off redux" results page on github. [9, 163, 165]. The top performing algorithms varied by dataset:

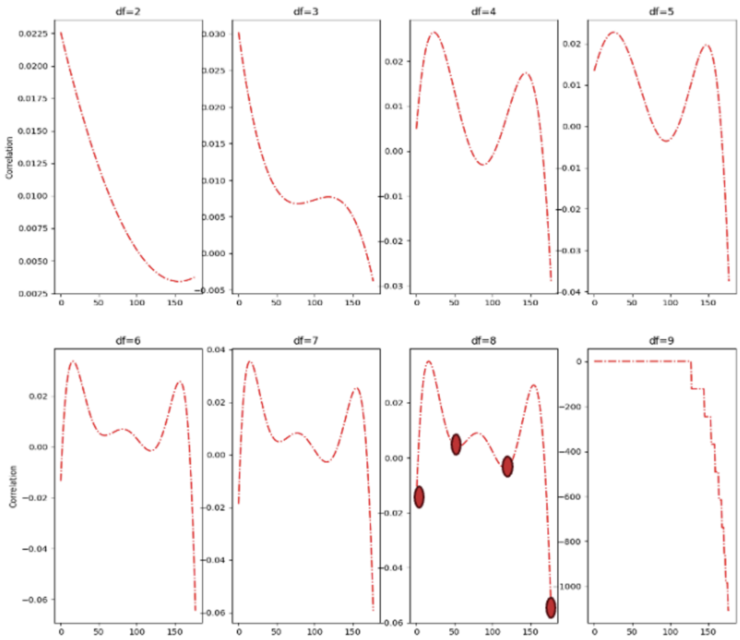

Figure 3.2: Polynomial Fitting for the Epilepsy Dataset

- Epilepsy: HIVE-COTE (Hierarchical Vote Collective of Transformation Ensembles) [154]
- SelfRegulationSCP1: TapNet [166]
- Earthquakes: Signatures [167]
- ItalyPowerDemand: Randomised Supervised Time Series Forest (RSTSF) [142]
- SharePriceIncrease: FRESHPrince (Fresh Pipeline with RotatIoN forest Classifier) [122]

HIVE-COTE and Hydra(Hybrid Dictionary-based Rocket)+MultiROCKET represent the overall top performing algorithms across time-series classification datasets in the redux and the original bake off paper [9, 127].

### 3.4.5 Results

#### Classification Accuracy

Table 3.1 presents the top 5 scores for each algorithm (out of 30 benchmark algorithms) for classification accuracy across all datasets. Results were taken from the TSML-EVAL github page [165]. Results represent average accuracy from 30 independent runs.

| **Dataset** | **Top 5 Results (Average Accuracy)** | | | | | **OTS-IVFS** |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | |
| Epilepsy | **1.000** | 0.999 | 0.999 | 0.998 | 0.998 | 0.921 |
| SelfRegulationSCP1 | **0.957** | 0.879 | 0.870 | 0.867 | 0.866 | 0.824 |
| Earthquakes | 0.767 | 0.756 | 0.752 | 0.751 | 0.750 | **0.895** |
| ItalyPowerDemand | **0.967** | 0.965 | 0.965 | 0.964 | 0.964 | 0.940 |
| SharePriceIncrease | **0.694** | 0.691 | 0.690 | 0.686 | 0.686 | 0.686 |

Table 3.1: Comparison of top 5 average accuracies and OTS-IVFS results for different datasets

OTS-IVFS achieved substantial improvement on the Earthquakes dataset with an accuracy of 0.895 compared to the best baseline of 0.767, representing a 16.67% increase in accuracy. Performance on the SharePriceIncrease dataset matched the fourth and fifth-ranked algorithm (0.686). The ItalyPowerDemand and Epilepsy datasets showed competitive performance with accuracies of 0.940 and 0.921 respectively, though below the top-performing algorithms. The SelfRegulationSCP1 dataset demonstrated reduced performance (0.824) compared to the leading TapNet baseline (0.957).

### Feature Reduction

The optimization process in OTS-IVFS substantially reduces the dimensionality of the feature space, as demonstrated in Table 3.2, which compares feature counts between traditional FRBS and our proposed method. Notably, benchmark algorithms in time-series classification often rely on hundreds or even thousands of features. For instance, the FreshPRINCE approach extracts approximately 794 features for model construction [122]. In contrast, the significantly smaller number of features used by OTS-IVFS represents a major advancement in model interpretability without compromising competitive performance.

Table 3.2: Feature Space Comparison

| Dataset | Traditional FRBS | OTS-IVFS |
|---|---|---|
| Epilepsy | 65 | 21 |
| Earthquakes | 120 | 18 |
| SharePriceIncrease | 45 | 15 |

**Rule Interpretability**

The generated rules demonstrated enhanced interpretability through linguistic time periods. Example rules from the Epilepsy classification included:

For Class 0 (No Epilepsy):

- IF Around T-60 is medium AND Around T-174 is medium THEN class is 0
- IF Around T-0 is medium AND Fixed Diff Around T-10 is medium THEN class is 0

For Class 1 (Epilepsy):

- IF Around T-0 is high AND Fixed Diff Around T-10 is high THEN class is 1
- IF Around T-94 is low AND Fixed Diff Around T-10 is low THEN class is 1

The rules reveal patterns where medium values indicate non-epileptic cases whilst extreme values suggest epileptic seizures. The ease of understanding patterns from these time specific rules and yet incorporating fuzziness around time-steps (such as Around T-94) underlines the key objective of increasing explainability with our proposed method.

**Comparison with Traditional FRBS**

A traditional IT2FRBS without time-based fuzzification generated rules such as:

- IF T-20 is high THEN class is 0
- IF T-70 is low AND T-71 is low AND T-72 is low THEN class is 0

These rules lack the semantic clarity of time-based fuzzy sets. The clustering of rules around T-70 to T-73 could be more effectively expressed through a single time-based fuzzy set.

### 3.4.6 Discussion

The experimental results validate the OTS-IVFS approach for time-series classification. The system maintained competitive accuracy whilst substantially improving model interpretability. The Earthquakes dataset demonstrated that OTS-IVFS can surpass state-of-the-art methods in specific domains.

Feature reduction through time-step selection enhanced computational efficiency without sacrificing performance. The linguistic representation of time periods in rules provides domain experts with an intuitive understanding of classification decisions. The algorithms with the highest performance on each dataset are: HIVE-COTE, TapNet, Signatures, RSTSF and FreshPRINCE. These algorithms employ opaque classification methodologies. RSTSF constructs an ensemble of binary classifiers using randomised intervals. TapNet operates through feature extraction via a random group permutation method integrated with multi-layer convolutional networks. Signatures constitute a feature-based approach that employs collections of ordered cross-moments [9, 166, 167]. FreshPrince builds a Rotation Forest model on around 800 features extracted from the time-series [122].

The results indicate that for the Earthquakes dataset, OTS-IVFS achieved superior performance relative to other algorithms. For the SharePriceIncrease, ItalyPowerDemand, and Epilepsy datasets, OTS-IVFS performance approximated that of the top five algorithms. OTS-IVFS demonstrated reduced performance on the SCP1 dataset compared to the top five algorithms.

According to [9], two algorithms have demonstrated consistent performance across time-series classification datasets: HIVE-COTE and Hydra+MultiROCKET, the latter based on the ROCKET approach. HIVE-COTE represents an ensemble method combining modules from distance-based representations, interval-based methods, dictionary-based approaches, shapelet-based techniques, and spectral-based methods. The updated version incorporates inputs from the ROCKET classifier family. ROCKET functions as a pipeline classifier applying a ridge classifier to features generated from numerous randomly parameterised convolutional kernels. Data transformation occurs through pooling operations including maximum and proportion of positive values. ROCKET-based approaches typically generate feature sets containing thousands of features [9].

Both HIVE-COTE and ROCKET-based approaches exhibit complexity and opacity due to their model structure and feature characteristics. OTS-IVFS employs interpretable features and transparent feature processing, facilitating inference and providing classification insights.

The results support the research objective that OTS-IVFS identifies and employs relevant time periods, thereby improving fuzzy time set construction and classification explainability while maintaining performance. The findings suggest that the presented approach can enhance explainability while delivering competitive classification performance in certain applications.

## 3.5 Case Study on ECG 200 Dataset

In this case study, we apply the OTS-IVFS methodology to the *ECG200* dataset from the UCR Time Series Classification Archive. The objective of this study is to demonstrate how the OTS-IVFS approach can be applied to biomedical time-series data to produce an interpretable classification model. Specifically, we illustrate the process of selecting predictive time steps, constructing time-based fuzzy sets, and generating interval type-2 fuzzy rules for classification. The results highlight how the fuzzy rule-based framework can provide insights into the temporal patterns associated with different ECG classes while maintaining competitive predictive performance.

This case study therefore contributes an applied demonstration of explainable time-series classification using fuzzy logic, illustrating the practical benefits of combining temporal feature selection with interval type-2 fuzzy rule-based modeling in the context of ECG signal analysis.

### 3.5.1 Dataset Description

The experiments in this case study are conducted using the *ECG200* dataset from the UCR Time Series Classification Archive, a widely used benchmark repository for evaluating time-series classification algorithms. The dataset contains electrocardiogram (ECG) signals representing two distinct classes corresponding to normal and abnormal heartbeats.

Each instance in the ECG200 dataset consists of a univariate time series representing a single heartbeat segment recorded from an ECG signal. Every time series contains 96 time steps, capturing the temporal progression of electrical activity during a cardiac cycle. The dataset comprises a total of 200 labeled instances, with 100 instances belonging to each class. Following the standard split defined in the UCR archive, the dataset is divided into 100 training samples and 100 testing samples.

The two classes in the dataset represent different cardiac conditions:

- **Class 1:** Normal ECG waveform.

- **Class 2:** Abnormal ECG waveform, often associated with myocardial infarction.

ECG signals typically contain characteristic waveform components such as the P-wave, QRS complex, and T-wave, which correspond to different phases of cardiac electrical activity. Differences in the morphology and timing of these components can indicate abnormalities in cardiac function. As a result, ECG data provides a meaningful real-world example for evaluating time-series classification methods.

One of the key challenges in analyzing ECG time series lies in capturing the temporal dependencies that occur at specific points in the cardiac cycle. Small variations in amplitude or timing may correspond to clinically significant differences. Traditional classification methods often treat each time step independently or rely on feature extraction techniques that summarize the entire signal. However, such approaches may overlook important localized patterns in the time series.

The ECG200 dataset is therefore well suited for evaluating the proposed OTS-IVFS methodology, as the approach explicitly identifies the most predictive time steps and constructs fuzzy sets around these temporal regions. By aggregating information from neighbouring time steps and representing them through fuzzy membership functions, the model can capture important waveform characteristics while maintaining interpretability [168].

Table 3.3 summarizes the key characteristics of the dataset.

Table 3.3: Summary of the ECG200 Dataset

| **Characteristic** | **Value** |
|---|---|
| Dataset Name | ECG200 |
| Source | UCR Time Series Classification Archive |
| Total Instances | 200 |
| Training Instances | 100 |
| Testing Instances | 100 |
| Number of Classes | 2 |
| Time Steps per Instance | 96 |
| Data Type | Univariate Time Series |
| Application Domain | Biomedical / ECG Signals |

The relatively small size of the ECG200 dataset makes it particularly suitable for interpretable models such as fuzzy rule-based systems. In the following section, we describe how the OTS-IVFS framework is applied to this dataset, including the process of predictive time-step selection, fuzzy set construction, and rule-based classification.

### 3.5.2 Methodology: Application of OTS-IVFS to ECG200

This section describes how the Optimized Time Series Interval Valed Fuzzy System (OTS-IVFS) framework is applied to the ECG200 dataset for time-series classification. The methodology follows the three-stage approach proposed in the original work, consisting of (1) predictive time-step selection, (2) time-based fuzzy set construction, and (3) training an Interval Type-2 Fuzzy Rule-Based System (IT2FRBS) for classification [169].

The goal of this approach is to identify the most informative temporal regions of the ECG signal and construct interpretable fuzzy rules that explain the classification decision.

**Overview of the OTS-IVFS Framework**

As discussed in previous sections, the OTS-IVFS framework is designed to address two key challenges in time-series classification: identifying relevant temporal features and maintaining model interpretability. Unlike traditional approaches that treat each time step as an independent feature, OTS-IVFS groups neighbouring time steps into fuzzy temporal regions, capturing local temporal patterns while reducing sensitivity to noise.

The methodology consists of the following steps:

1. Selection of predictive time steps from the ECG time series.
2. Construction of optimal time-based fuzzy sets around the selected time steps.
3. Generation of fuzzy features and training of an Interval Type-2 Fuzzy Rule-Based System for classification.

Each of these stages is described in detail below.

**Predictive Time-Step Selection**

The first stage identifies the time steps that have the strongest relationship with the class labels in the ECG200 dataset. Rather than using all 96 time steps as independent input variables, the algorithm determines which regions of the ECG waveform contain the most discriminative information.

The selection process begins by computing the correlation between the signal value at each time step and the class label across the training dataset. This produces a correlation time series representing the predictive importance of each time index as shown in Figure 3.3.

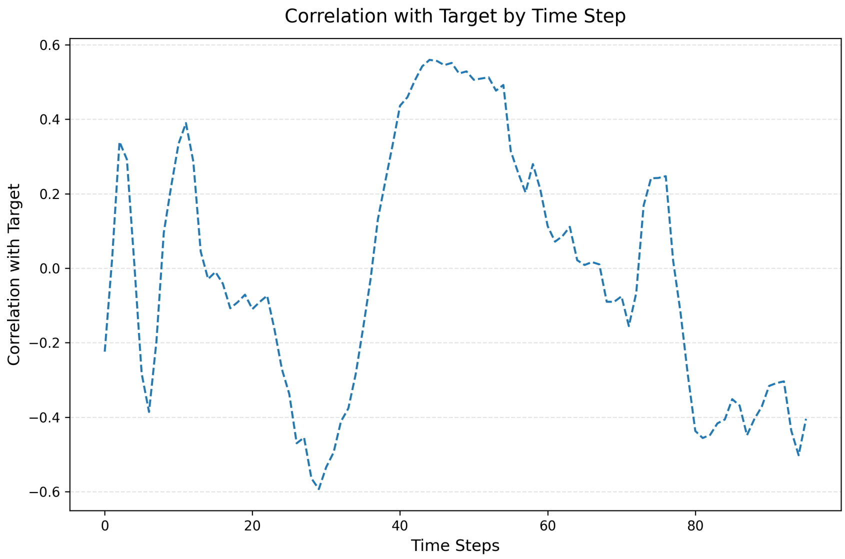


Figure 3.3: Correlations of Time Steps with Target on ECG 200 Dataset

Following this, polynomial functions of increasing degree are fitted to the correlation curve. The purpose of this step is to identify significant local maxima and minima that indicate meaningful temporal regions in the ECG signal. The selected time steps correspond to peaks in the fitted curve that represent the strongest positive or negative correlations with the target class. The fits for correlation are as shown in Figure 3.4. Similarly, the process is run for negative correlations, since the correlation time-series suggests that there are significant negative correlations as can be clearly seen in Figure 3.3. This process allows the algorithm to automatically determine the number and location of relevant time steps without requiring manual specification. By selecting well-separated time points, the model avoids redundant features and focuses on the most informative parts of the signal [169]. The combined time-steps selected at this stage were: [2, 6 11, 29, 44, 71, 76, 92, 94].

### Time-Based Fuzzy Set Construction

Once the predictive time steps have been identified, fuzzy sets are constructed around these time points to capture local temporal patterns. Instead of using a single time step as a feature, the algorithm defines a fuzzy membership function over a small temporal window surrounding the selected time index.

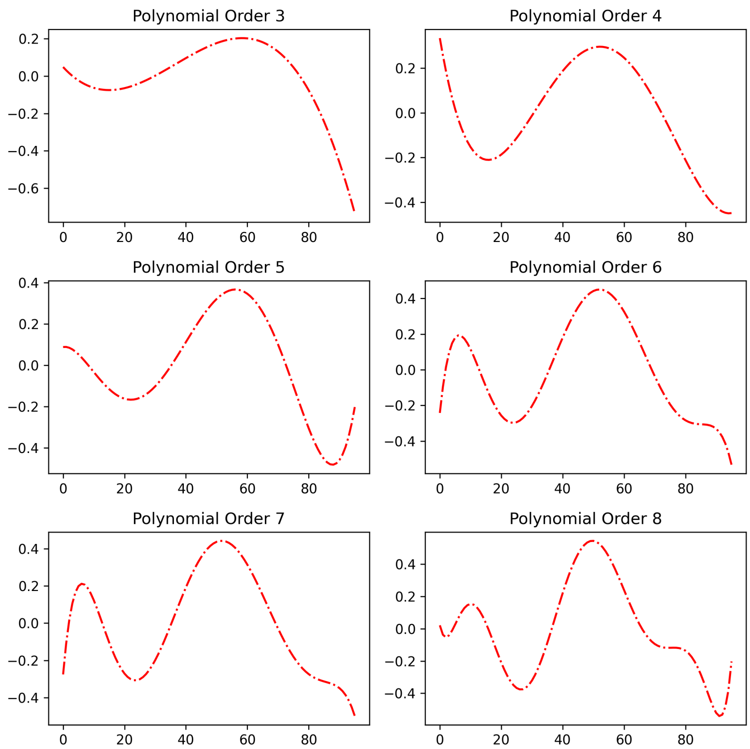


Figure 3.4: Polynomial Fitting for Time Step Selection on ECG 200 Dataset

For each selected time step $t$, a range $[t-w, t+w]$ is defined, where $w$ represents the maximum allowable width of the fuzzy set. Within this range, multiple candidate fuzzy sets are generated using common membership function shapes such as triangular, trapezoidal, and shoulder functions.

Each candidate fuzzy set is applied to the training data to transform the original ECG values into fuzzy membership values. The quality of each fuzzy set is evaluated using a performance metric such as ROC-AUC or classification accuracy. The fuzzy set that yields the best performance is selected as the optimal representation for that time region.

This step effectively converts raw ECG signals into interpretable temporal features that describe signal behaviour around clinically relevant points in the waveform.

### Generation of Fuzzy Features

After determining the optimal fuzzy sets, the ECG time series is transformed into a new feature space. Each fuzzy set produces a feature representing the degree to which the signal values around a selected time step belong to the corresponding fuzzy region.

For example, if a fuzzy set is constructed around time step $t$, the resulting feature may represent the aggregated membership of signal values within the interval $[t-w, t+w]$. These features capture local signal behaviour while smoothing out small fluctuations or noise.

Additional derived features such as first-order differences and fixed differences between time steps can also be included to capture dynamic changes in the ECG waveform.

### Training the Interval Type-2 Fuzzy Rule-Based System

The final stage of the methodology involves training an Interval Type-2 Fuzzy Rule-Based System using the extracted fuzzy features.

The training process consists of the following steps:

1. **Fuzzification:** Input features are converted into linguistic fuzzy variables such as *Low*, *Medium*, and *High*.

2. **Rule Generation:** A set of fuzzy IF–THEN rules is constructed using combinations of the fuzzy variables.

3. **Rule Optimization:** Optimization algorithms are used to select an optimal subset of rules that maximizes classification accuracy while maintaining interpretability.

4. **Inference:** For a new ECG signal, the classifier evaluates the activated rules and predicts the class with the highest vote strength.

### Advantages for ECG Classification

Applying OTS-IVFS to ECG200 offers several benefits:

- **Interpretability:** The model produces human-readable rules that explain which regions of the ECG signal influence the classification.

- **Noise Robustness:** Fuzzy sets aggregate information across neighbouring time steps, reducing sensitivity to noise in the signal.

- **Feature Reduction:** Only a small number of informative time steps are selected, reducing dimensionality compared to using all 96 time points. In this case, 9 time-steps were selected as being the most informative, representing a substantial drop in the input feature space.

These properties make the approach particularly suitable for biomedical applications, where both accuracy and interpretability are essential for supporting clinical decision-making.

### 3.5.3 Experiments and Results

To evaluate the effect of rule complexity on model performance, experiments were conducted using rule bases with maximum rule lengths ranging from 1 to 4 antecedents. For each configuration, the model was trained 30 times and evaluated using three metrics on the out of sample dataset as provided in the UCR archive: Accuracy, F1-score, ROC_AUC. The total number of rules generated in the final rule base was also recorded as a measure of complexity of the model.

Figures 3.5, 3.6 and 3.7 illustrate the distribution of the evaluation metrics across the different rule lengths.

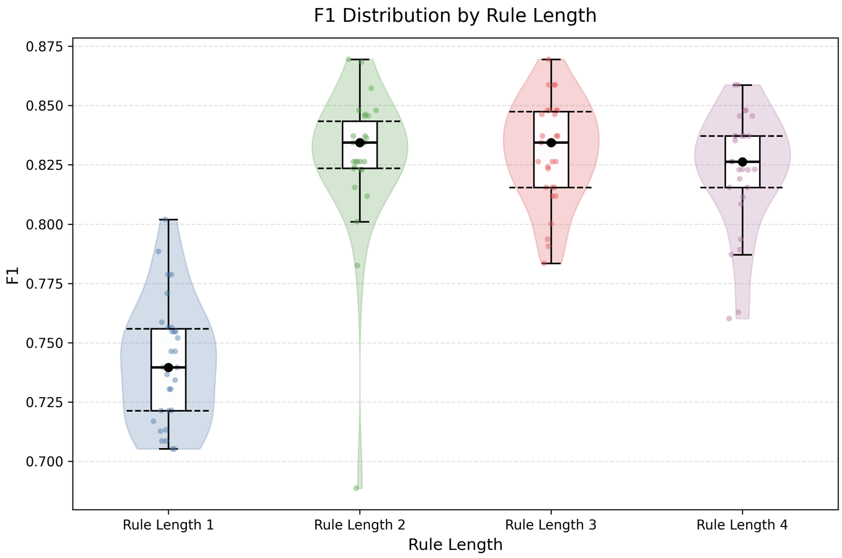


Figure 3.5: F1 Score Distributions (30 runs) by Rule Length on ECG 200 Test Dataset

**Accuracy and F1-score** The results show a clear improvement in performance when moving from rule length 1 to rule length 2. Models restricted to single-antecedent rules achieved noticeably lower performance, with accuracy and F1 scores centered around approximately 0.75. This indicates that rules containing only one temporal feature are insufficient to capture the relationships within the ECG signals.

Allowing rules with two antecedents significantly improves predictive performance. For rule length 2, both accuracy and F1 scores increase to 0.84–0.86 range. This suggests that

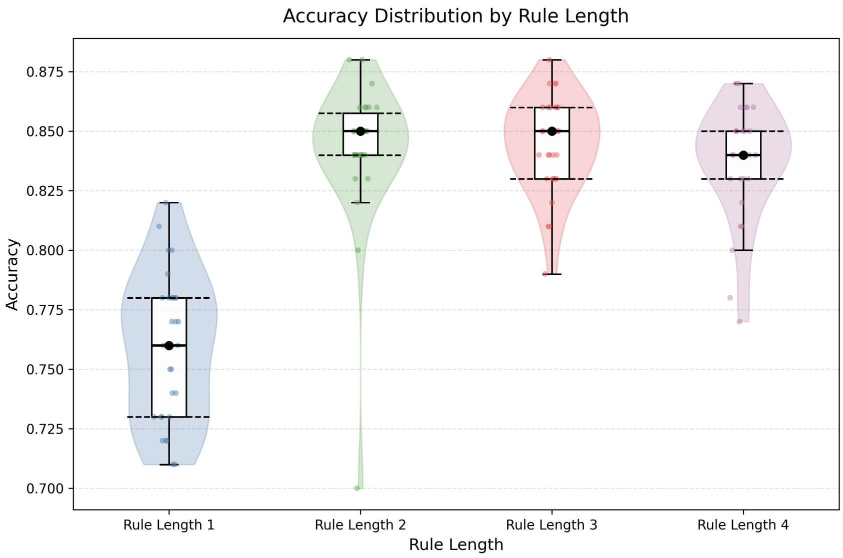


Figure 3.6: Accuracy Distributions (30 runs) by Rule Length on ECG 200 Test Dataset

combining information from multiple temporal regions of the ECG signal enables the fuzzy rule-based system to better distinguish between the two classes.

Increasing the rule length further to three antecedents yields only marginal improvements. The distributions of both accuracy and F1-score for rule length 3 remain similar to those observed for rule length 2, indicating diminishing returns from additional rule complexity. When rule length is extended to four antecedents, the performance deteriorates slightly, potentially pointing to the optimization algorithm struggling to find an optimmal rule-base given a bigger underlying universe of rules.

**ROC_AUC Performance** A similar trend can be observed in the ROC_AUC results. The model with rule length 1 achieves ROC_AUC values centered around 0.85. When rule length increases to 2, the ROC_AUC improves significantly to around 0.93.

Further increases in rule length do not substantially improve this metric. Rule lengths 3 and 4 achieve ROC_AUC values that remain within a similar range, suggesting that the model already captures most of the discriminative temporal information when using two or three antecedents per rule. However, the rule length 4 does show the highest average ROC_AUC score of all options, suggesting that the classification performance is threshold dependent (as F1 and accuracy scores are both recorded at a single threshold of 0.5).

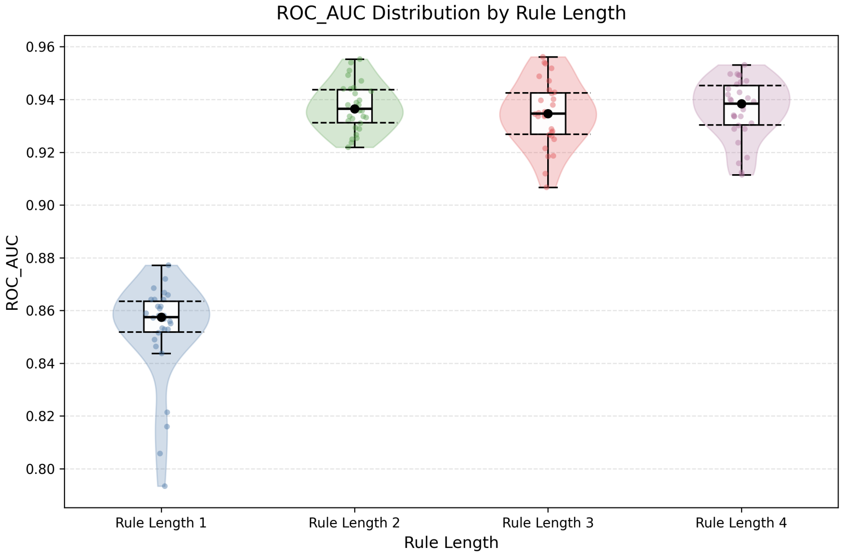


Figure 3.7: ROC_AUC Distributions (30 runs) by Rule Length on ECG 200 Test Dataset

**Model Complexity and Number of Rules** While predictive performance stabilizes after rule length 2, model complexity increases substantially as rule length grows. As shown in Figure 3.8, the number of rules in the optimized rule base grows rapidly with rule length.

For rule length 1, the rule base contains 10–20 rules on average. Increasing the rule length to 2 leads to a substantial expansion, with the rule base containing roughly 90–110 rules. When rule length reaches 3, the number of rules increases further to approximately 170–190. Rule length 4 produces similarly large rule bases, often exceeding 160 rules but with more variation in the rule base size which could again point to the optimization algorithm producing more variable rule bases when the underlying universe of rules becomes larger.

This rapid growth in the number of rules highlights the trade-off between interpretability and model complexity. Larger rule bases make the model more difficult to interpret and may introduce redundancy without improving predictive performance.

**Interpretability vs Performance Trade-off** The experimental results demonstrate that rule length 2 provides the best balance between predictive performance and model interpretability. This configuration achieves nearly the same accuracy, F1-score, and ROC_AUC as more complex rule structures while maintaining a substantially smaller and more interpretable rule base.

Consequently, restricting the maximum rule length to two antecedents allows the OTS-IVFS model to capture meaningful temporal relationships in the ECG signals without un-

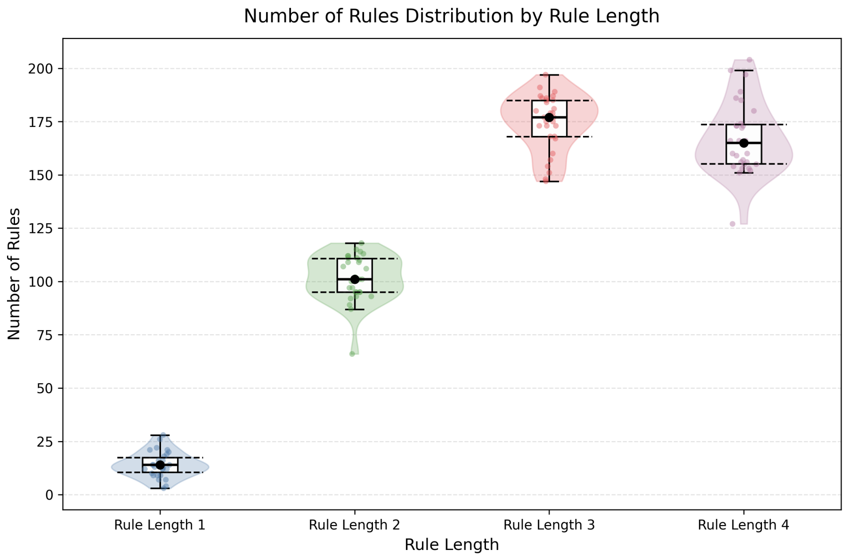


Figure 3.8: Number of Rules Distributions (30 runs) by Rule Length on ECG 200 Test Dataset

necessarily increasing the complexity of the rule base.

These findings align with the objective of the OTS-IVFS framework, which aims to produce accurate yet interpretable classification models for time-series data.

### 3.5.4 Comparison with Published ECG200 Results

To contextualize the performance of the proposed OTS-IVFS model, it is useful to compare the results obtained on the ECG200 dataset against representative results reported for other time-series classification approaches evaluated on the same benchmark. The ECG200 dataset in the UCR archive contains 100 training instances and 100 testing instances, with each series having length 96 and the two classes corresponding to normal heartbeat and myocardial infarction. In this case study, we follow the fixed out-of-sample train/test partition provided by the UCR archive.

Table 3.4 summarizes a set of comparable published results. In the ROCKET comparison table, ECG200 accuracies reported on the archive split include 0.906 for ROCKET, 0.909 for Proximity Forest, 0.870 for BOSS, 0.830 for Shapelet Transform, 0.874 for ResNet, 0.850 for HIVE-COTE, 0.910 for InceptionTime, and 0.860 for TS-CHIEF [143]. A separate InceptionFCN study reports 0.94 for InceptionFCN and 0.93 for InceptionTime on ECG200 [170]. It is worth noting that the higher accuracy scores in InceptionFCN study were not on the default train-test split but were median results on a number of randomised splits and are hence not directly comparable. However, the results are included for completeness.

Table 3.4: Comparison with published ECG200 test accuracy results

| Method | Accuracy | Interpretability |
|---|---|---|
| OTS-IVFS (this study, best range) | $\approx$ 0.85 | High |
| Shapelet Transform | 0.830 | Moderate |
| TS-CHIEF | 0.860 | Low |
| BOSS | 0.870 | Low |
| ResNet | 0.874 | Low |
| ROCKET | 0.906 | Low |
| Proximity Forest | 0.909 | Low |
| InceptionTime | 0.910–0.930 | Low |
| InceptionFCN | 0.940 | Low |

The comparison shows that the proposed OTS-IVFS does not match the very best black-box classifiers on raw test accuracy, but the gap is relatively modest given the substantial gain in explainability. Based on the distributions in the previous subsection, the strongest OTS-IVFS configurations achieve an out-of-sample accuracy of about 0.85, with F1 values in a very similar range and ROC_AUC close to 0.94. This places the method close to published results for Shapelet Transform (0.83), TS-CHIEF (0.86), BOSS (0.87), and ResNet (0.874), while remaining below ROCKET and InceptionTime-style convolutional approaches.

From an applied perspective, this is an important result: the price paid for interpretability is only around one to a few percentage points relative to several established baselines, and roughly five to nine percentage points relative to the strongest deep learning and convolution-kernel methods reported here. In return, OTS-IVFS yields an explicit fuzzy rule base, readable linguistic labels, and direct temporal explanations in terms of informative regions of the ECG waveform, rather than relying on opaque latent features or large ensembles. This trade-off is especially relevant in medical settings, where clinicians may value transparency and traceability alongside predictive performance.

It is also worth noting that ECG200 is known to have benchmark-specific quirks. The bake-off study highlights that ECG200 was incorrectly formatted in a way that can bias results, and notes that the sum of squares of the series can classify the data perfectly if a classifier exploits that artifact. This means that small differences in leaderboard accuracy on ECG200 should be interpreted cautiously, especially when comparing highly optimized black-box models. Under this caveat, the fact that OTS-IVFS achieves about 0.85 accuracy on the fixed out-of-sample UCR split while remaining inherently explainable is a strong result.

Overall, the comparison supports the central claim of this case study: OTS-IVFS preserves competitive predictive performance on ECG200 while delivering a much higher level of interpretability than most published alternatives. Rather than optimizing solely for leaderboard performance, the method offers a balanced solution in which the classifier remains transparent, its rules can be inspected directly, and the temporal basis of each decision can be explained in human-readable terms.

## 3.6 Limitations and Critical Analysis

This section discusses some limitations of the proposed approach followed by a discussion on the comparison with the state-of-the-art methods and the key contributions.

### 3.6.1 Limitations

- **Scalability with High-Dimensional Time Series** The OTS-IVFS approach exhibits limitations when processing time series with extensive temporal dimensions. The experimental results demonstrate reduced performance on the SelfRegulationSCP1 dataset, which contains 896 time-steps. This degradation suggests that the polynomial-based time-step selection may struggle to identify meaningful patterns in high-dimensional temporal spaces. Similar challenges with long time series have been documented in fuzzy time-series literature [117].

  The correlation-based approach assumes that predictive patterns can be captured through polynomial fitting. However, for time series exceeding several hundred observations, this assumption may oversimplify complex temporal dependencies. A potential limitation of OTS-IVFS is that the correlation-based time-step selection process may overlook locally informative temporal patterns that do not exhibit strong global correlation with the target variable. In such cases, a time-step may capture a predictive phenomenon that is conditionally important or relevant only in specific contexts, and therefore may not be selected despite its potential contribution when combined with other features.

- **Computational Complexity** The computational burden of OTS-IVFS increases substantially with problem dimensions. The time-step selection algorithm requires $O(n \cdot d_{max}^2 \cdot K)$ operations where $n$ represents the look-back window, $d_{max}$ denotes the maximum polynomial degree, and $K$ indicates the number of input time series. The fuzzy set optimization phase compounds this complexity with $O(|T_{selected}| \cdot S \cdot w_{max} \cdot N \cdot n)$ evaluations, where $|T_{selected}|$ represents selected time-steps, $S$ denotes fuzzy set shapes tested, $w_{max}$ indicates maximum width, $N$ represents training instances, and $n$ denotes look-back window size. The genetic algorithm optimization adds another layer of computational expense with complexity $O(G_{max} \cdot |P| \cdot |RB|^2 \cdot N)$ where $G_{max}$ represents generations, $|P|$ denotes population size, and $|RB|$ indicates rule base size. With 100 chromosomes evolved over 50 generations, each requiring full model evaluation, the training time becomes substantial compared to methods like Random Forest.

- **Univariate Time Step Selection** A fundamental limitation lies in the time-step selection methodology. The current approach identifies predictive time-steps based solely on the primary time-series correlation with the target variable. This uni-variate perspective fails to accommodate multivariate scenarios where different input series exhibit varying predictive ranges.

  Consider a financial forecasting scenario with multiple economic indicators. Interest rates might show predictive power at 3-month intervals, whilst employment data proves most informative at 1-month lags. The current methodology would impose uniform time-steps across all series, potentially missing these variable temporal relationships.

  This limitation becomes particularly acute in heterogeneous sensor networks or multi-modal data fusion applications. Different sensors may operate at different sampling rates or exhibit distinct temporal characteristics. The uniform time-step selection cannot capture these nuanced temporal patterns [127].

- **Limited Handling of Non-Stationary Patterns** The correlation-based time-step selection assumes relatively stable temporal relationships. However, many real-world time series exhibit non-stationary behaviour with evolving patterns over time. Financial markets demonstrate regime changes, sensor networks experience drift, and biological signals show adaptive responses. The static time-step selection cannot accommodate these dynamic temporal characteristics.

The polynomial fitting approach further assumes smooth correlation patterns amenable to low-degree polynomial approximation. Time series with abrupt changes, seasonal variations, or multiple periodicities may not conform to this assumption. Recent work on adaptive fuzzy time-series has highlighted the importance of dynamic model adjustment [31].

The interpretability advantage diminishes as rule bases expand. While individual rules remain comprehensible, understanding the collective behaviour of hundreds of rules challenges human cognitive capabilities. This issue mirrors problems identified in traditional fuzzy systems literature [98].

- **Limited Cross-Domain Generalization** The experimental results reveal inconsistent performance across domains. While OTS-IVFS excelled on seismic data (Earthquakes dataset), it underperformed on brain-computer interface data (SelfRegulationSCP1). This variability suggests that the approach may require domain-specific tuning, limiting its generalisation as a universal time-series classification method.

  The time-based fuzzy sets, whilst interpretable, impose specific assumptions about temporal locality. Domains with long-range dependencies or complex temporal interactions may not align with these assumptions. Deep learning approaches like LSTMs can capture such patterns more flexibly, albeit without interpretability [5].

### 3.6.2 Comparison with State-of-the-Art Methods

Evaluating OTS-IVFS against current state-of-the-art methods reveals both strengths and limitations. The approach cannot match the raw performance of ensemble methods like HIVE-COTE or deep learning architectures like Hydra+MultiROCKET on most datasets. These methods achieve superior accuracy through complex model combinations or thousands of convolutional features [9].

However, positioning OTS-IVFS purely on accuracy metrics misses its primary contribution. The system occupies a unique position in the accuracy-interpretability trade-off space. While not achieving state-of-the-art accuracy, it provides genuinely interpretable models with competitive performance.

### 3.6.3 Significance of Contribution

Despite the limitations, OTS-IVFS represents a significant contribution to explainable time-series classification. The approach addresses a critical gap between high-performing opaque models and interpretable but simplistic methods. The linguistic representation of temporal patterns through optimized fuzzy sets offers a novel perspective on time-series interpretability.

The method's significance lies not in surpassing accuracy benchmarks but in demonstrating that meaningful interpretability can coexist with reasonable performance. For domains requiring explainable decisions—medical diagnosis, financial risk assessment, industrial fault detection—this trade-off proves valuable.

Future research can address these limitations through adaptive time-step selection for multivariate series, efficient optimization strategies for high-dimensional data, and dynamic adjustment mechanisms for non-stationary patterns. Extensions incorporating domain knowledge into time-step selection could enhance both performance and interpretability.

The OTS-IVFS contribution should be evaluated within the broader context of explainable AI research. As regulatory requirements and ethical considerations increasingly demand transparent AI systems, approaches balancing performance with genuine interpretability become essential. In this context, OTS-IVFS offers a meaningful step toward practical explainable time-series classification.

# Summary

This chapter addresses the growing need for explainable time-series classification methods by introducing the Optimized Time Series Interval Type-2 Fuzzy System (OTS-IVFS). Motivated by the limitations of both black-box models and existing fuzzy time-series approaches, the chapter reviews the evolution of time-series classification, highlighting gaps in temporal abstraction, linguistic interpretability, and noise robustness.

To overcome these challenges, the chapter proposes a fully data-driven and interpretable framework that automatically discovers informative time periods, optimizes fuzzy time sets, and constructs an interval type-2 fuzzy rule-based classifier. By aggregating time-steps into linguistically meaningful temporal features, OTS-IVFS aligns model representations with human temporal reasoning while improving robustness and reducing dimensionality. Experimental results across multiple benchmark datasets demonstrate that the proposed method achieves competitive accuracy with state-of-the-art black-box models, while producing transparent, human-readable rules. Overall, the chapter establishes OTS-IVFS as a practical and explainable alternative for time-series classification in domains where trust, validation, and interpretability are essential.

# Chapter 4

# Addressing Logical Gaps in Fuzzy Logic Systems for Regression

## 4.1 Introduction

In this chapter, we present a focused study of *logical gaps* in fuzzy rule-based regression systems (FRBSs), with particular emphasis on Mamdani-type models. FRBSs are widely valued for their transparent IF–THEN rule structure and their suitability for applications that require human interpretability. However, this same transparency exposes situations in which the rule base fails to support predictions that are consistent with accepted directional or monotonic relationships between inputs and outputs. The chapter therefore moves beyond questions of interpretability to address the internal logical soundness of rule-based reasoning.

We define a logical gap as a subset of the input space in which the rule base produces outputs that violate an expected directional relation. Unlike random prediction errors, logical gaps reflect structural inconsistencies in the rule base and persist even in the absence of noise. In such regions, small changes in inputs may lead to implausible jumps, reversals, or plateaus in the predicted output. These inconsistencies undermine user trust, complicate governance and compliance processes, and may trigger manual intervention that offsets the operational benefits of automated decision support.

Logical gaps arise naturally in practical deployments of FRBSs. Historical datasets are often unevenly distributed, with dense coverage in common operating regions and sparse observations near boundaries or under stress conditions. Rule induction and optimisation procedures, which typically prioritise global performance metrics such as RMSE or accuracy, tend to favour compact rule bases and may prune rules that are essential for maintaining consistent directional behaviour. Subsequent optimisation of membership functions or rule merging can further reduce model complexity, but does not guarantee preservation of monotonic trends and may even exacerbate inconsistencies by extending fuzzy sets beyond their empirical support.

A common response to sparse rule coverage is to employ fuzzy rule interpolation (FRI), which enables inference in uncovered regions by interpolating between neighbouring rules. While FRI has proven effective for improving numerical coverage and continuity, it does not fundamentally resolve logical gaps. Interpolation mechanisms are inherently derivative: interpolated conclusions are constructed as combinations or transformations of existing rules and therefore inherit the directional properties of the surrounding rule base. When neighbouring rules encode inconsistent or misaligned consequents, interpolation merely smooths between them without enforcing global monotonicity or repairing broken directional links. Even advanced interpolation schemes that incorporate rule weighting, ranking, or neighbourhood selection remain reliant on the semantic coherence of the original rules and cannot introduce missing directional knowledge where it is structurally absent. As a result, logical gaps caused by missing or contradictory rule consequents persist, albeit in a softened form.

Manual correction of such inconsistencies is possible, but becomes impractical as dimensionality increases and is vulnerable to omissions and subjective bias. Moreover, optimisation-driven interpolation and approximation methods may achieve low prediction error while still violating domain-level constraints, particularly in low-density regions of the input space where directional behaviour is most critical.

To address these limitations, this chapter proposes a targeted detection and repair method for logical gaps. The method first identifies input variables that exhibit strong directional relationships with the output. It then evaluates the rule base using a structured mapping between antecedent regions and output consequents to detect missing or inconsistent directional links. Where logical gaps are identified, new rules are inserted to restore the expected trend while preserving the surrounding validated rules and their parameters. When multiple candidate repairs are possible, a quality criterion is used to select the most appropriate rule. The approach is designed to correct specific inconsistencies without inflating the rule base or compromising interpretability.

The proposed method is evaluated using a real-world lending dataset, where monotonic relationships are expected between several risk-related variables and default outcomes. Performance is assessed in terms of both predictive accuracy and compliance with monotonicity constraints before and after gap repair. The results show that a small number of targeted rule insertions can close substantial logical gaps without material loss of accuracy and without significant growth in rule-base size.

The remainder of the chapter is organised as follows. Section 4.2 dives deeper into the rule interpolation mechanism and discusses why rule interpolation is not sufficient by itself to address logical gaps. Section 4.3 reviews the structure of Mamdani-type FRBS regression models, including fuzzification, rule bases, inference, and defuzzification. Section 4.4 formalises the notion of logical gaps, analyses their causes, and discusses their implications in domains governed by directional constraints. Sections 4.5 and 4.6 present the proposed detection and repair algorithms for single- and multiple-antecedent cases. This is followed by the experimental evaluation, and finally a discussion of conclusions, limitations, and directions for future research.

Through this sequence, the chapter demonstrates that interpretability alone is insufficient if internal logical relationships are not respected, and that targeted structural repair can restore coherence with minimal disruption to an existing rule base. The proposed technique supports the broader goal of deploying interpretable models that behave consistently with established domain knowledge.

## 4.2 Why Rule Interpolation Does Not Eliminate Logical Gaps

Fuzzy rule interpolation (FRI) has been widely studied as a means of enabling inference in regions of the input space that are not explicitly covered by existing rules. A substantial body of work demonstrates that interpolation can generate plausible outputs under sparse rule bases, thereby avoiding the need for exhaustive rule enumeration [171–173]. However, despite these advantages, rule interpolation does not fundamentally address the problem of logical gaps as defined in this chapter.

The primary limitation is that interpolation methods are inherently *derivative* rather than *corrective*. Interpolated conclusions are computed as weighted or transformed combinations of neighbouring rules and therefore inherit the structural properties—both desirable and undesirable—of the existing rule base. When adjacent rules encode inconsistent or weakly aligned directional relationships, interpolation merely smooths between them without enforcing global monotonicity or logical coherence. This issue is explicitly acknowledged in modern FRI variants, where improvements focus on numerical accuracy or local smoothness rather than on guaranteeing semantic consistency of trends [174–176].

More recent work has attempted to refine interpolation behaviour through rule weighting, ranking, or neighbourhood selection. For example, weighting strategies based on rule proximity or relevance [174, 177] and rule-ranking approaches incorporating directional monotonicity [178] can improve local inference quality. Nevertheless, these methods still assume that the underlying rule base provides a sufficiently coherent scaffold. If a monotonic relationship is structurally broken—because a necessary rule is missing or because consequents are misaligned—interpolation cannot introduce new directional information; it can only redistribute influence among existing rules. As a result, logical gaps caused by missing or contradictory rule consequents persist, albeit in a smoothed form.

A further limitation arises from the fact that many interpolation techniques are designed to operate under sparse but *consistent* knowledge assumptions. Group-based and evolutionary interpolation approaches for ANFIS construction [179,180] aim to approximate unknown regions while preserving overall predictive performance. However, their optimisation objectives are typically error-driven and do not explicitly encode domain-level constraints such as monotonicity. Consequently, an interpolated surface may achieve low prediction error while still violating expected directional behaviour, especially in boundary or low-density regions of the input space.

In summary, rule interpolation enhances coverage but does not repair structural deficiencies in a rule base. Logical gaps stem from missing or inconsistent directional links between antecedent regions and consequents, whereas interpolation operates by blending existing rules without questioning their semantic alignment. Closing logical gaps therefore requires explicit detection of violated directional relationships and targeted insertion or modification of rules, rather than reliance on interpolation alone. This distinction motivates the repair-based approach proposed in this chapter, which complements interpolation by restoring the internal logical soundness of fuzzy rule-based regression systems.

## 4.3 Mamdani fuzzy rule–based systems for regression

A Mamdani FRBS approximates a mapping $f : \mathbb{R}^n \to \mathbb{R}$. This structure was proposed by Mamdani and Assilian (1975) for control [29] and is now frequently applied to non-linear regression [181, 182]. The four main components of a Mamdani FRBS are: the *fuzzifier*, the *rule base*, the *inference engine* and the *defuzzifier* as shown in 2.1. Mamdani FRBSs have been discussed in detail in Section 2.2. Here, some relevant parts are discussed briefly to set the context.

### 4.3.1 Fuzzifier

Each crisp input $x_i$ is converted to membership degrees

$$\mu_{A_{ij}}(x_i) \in [0,1],$$

where $A_{ij}$ is the $j$-th linguistic term of variable $x_i$. Common membership functions are triangular, trapezoidal and gaussian. A singleton fuzzifier assigns $\mu_A(x) = 1$ at $x$ and 0 elsewhere, but continuous functions give smoother activation. The number and shape of the fuzzy sets is a subjective choice and is influenced by the problem type, data distribution, expert knowledge, and the desired levels of granularity and interpretability.

### 4.3.2 Rule base

The rule base contains $M$ rules of the form

$$R_k : \text{ IF } x_1 \text{ is } A_{1k} \text{ AND } \ldots \text{ AND } x_n \text{ is } A_{nk} \text{ THEN } y \text{ is } B_k, \qquad k = 1,\ldots,M. \tag{4.1}$$

Each antecedent set $A_{ik}$ partitions the input space and the consequent $B_k$ is a fuzzy set in the output domain. Rules may be provided by experts or induced from data by clustering, evolutionary search or gradient tuning [89].

The size of the potential universe of rules - i.e. rules available for building the model are influenced primarily by the number of input features, the number of fuzzy sets used to fuzzify the inputs and the limitations imposed on individual rules in terms of the maximum number of terms in the antecedent.

Evolutionary techniques have been a popular choice to optimize the FRBSs parameters, including the rule-base as discussed in Section 2.2.5. The learning approach to EFS focuses on rule learning and rule selection. As the number of inputs to an FRBS grows, the space of potential fuzzy rules becomes enormous - requiring intelligent search techniques to prune down to an optimal rule set - and it is in the pruning of the rules that the rule-bases sometimes become sparse and logical gaps are introduced.

### Rule quality

EAs usually need a measure to select a rule probabilistically. Towards this end, each rule can be assigned a measure of quality that combines two empirical properties: fuzzy support and fuzzy confidence. A simple product measure is then:

$$Q_l = S_l C_l, \tag{4.2}$$

where $Q_l$ is the quality of the $l$-th rule, $S_l$ its support and $C_l$ its confidence.

Let $n(\text{Ant}_l, \text{Cons}_l)$ denote the number of observations that satisfy both the antecedent $\text{Ant}_l$ and the consequent $\text{Cons}_l$. Let $n(\text{Ant}_l)$ be the number of observations that satisfy $\text{Ant}_l$. With $m$ the total sample size,

$$S_l = \frac{n(\text{Ant}_l, \text{Cons}_l)}{m}, \tag{4.3}$$

$$C_l = \frac{n(\text{Ant}_l, \text{Cons}_l)}{n(\text{Ant}_l)} \tag{4.4}$$

Where, support is a measure of how often the antecedent of the rule occurs in the data and confidence tells us that when the rule fires, how often does it fire to the correct consequent fuzzy set - thus giving us a measure for the accuracy of the rule. Any alternative measure that meets the needs of a particular prediction task may be substituted for $Q_l$. Alternatively, Confidence can be used as a simpler quality measure on its own, depending only upon the precision of the rule and not the coverage.

### 4.3.3 Inference engine

For an input $x$ each rule fires with strength

$$\alpha_k = T\big(\mu_{A_{1k}}(x_1), \ldots, \mu_{A_{nk}}(x_n)\big),$$

where $T$ is a $T$-norm, typically min or the product. In this study, the minimum operator is adopted because the product $T$-norm can rapidly drive firing strengths toward zero when several antecedent membership degrees are small. Using the minimum operator preserves the scale of the weakest antecedent without compounding small values, leading to more stable rule activation and a clearer correspondence between antecedent satisfaction and rule contribution. Implication scales the consequent:

$$\mu_{B_k^*}(y) = \alpha_k \circ \mu_{B_k}(y), \qquad \circ \in \{\text{min, product}\}.$$

Aggregated output is obtained by an $S$-norm, usually the maximum:

$$\mu_B(y) = \max_{k=1}^{M} \mu_{B_k^*}(y).$$

### 4.3.4 Defuzzifier

Although, there are many defuzzification techniques, one of the most popular techniques is the centroid defuzzification. The defuzzifier converts$\mu B(\cdot)$ to a crisp prediction $\hat{y}$. The centroid defuzzifier method combines the output sets using a t-norm as shown in 2.2 and 2.3.

## 4.4 Emergence of Logical Gaps in Fuzzy Rule-Based Systems for Regression

FRBS for regression are valued for their ability to produce interpretable and accurate models [63]. As discussed earlier, in data-driven FRBS design, rule bases are often constructed through optimisation algorithms such as EAs. The optimisation objective typically combines two goals: minimising prediction error and maintaining interpretability by limiting the number of rules and the number of antecedents per rule.

While effective, such optimisation strategies can result in rule bases that contain semantic inconsistencies, such as the presence of *logical gaps*. Formally, a logical gap arises when the following four conditions are satisfied:

1. There exists a known monotonic relationship between an input feature $x_i$ and the output feature $y$, either positive or negative. Such relationships encode basic directional knowledge about the system under study and may be commonly derived from physical laws, regulatory principles, or long-standing empirical evidence. Monotonicity therefore represents a minimal form of structural consistency that a predictive model is expected to satisfy. When this condition holds, any violation cannot be dismissed as acceptable variability but instead signals a breakdown in the internal coherence of the model.

2. An initial input vector

$$\mathbf{x} = ((x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_m)$$

   is passed through the FRBS to produce a predicted output value $o$.

3. A new input vector

$$\mathbf{x}^* = (x_1, \ldots, x_{i-1}, x_i^*, x_{i+1}, \ldots, x_m)$$

   is constructed by changing the value of $x_i$ to $x_i^*$, while keeping all other inputs constant. The new output is denoted as $o^*$.

4. The resulting change in output, $o^* - o$, is in a direction inconsistent with the known relationship between $x_i$ and $y$. For example, if the relationship between $x_i$ and $y$ is positive, then an increase in $x_i$ should produce an increase in $o^*$, i.e.,

$$x_i^* > x_i \Rightarrow o^* > o,$$

   but instead the inequality is reversed or violated.

This type of inconsistency indicates a logical gap in the rule base, where the FRBS fails to represent expected behaviour due to missing or insufficient rules. Focusing on missing or insufficient rules provides a clearer and more actionable repair strategy. Removing or modifying allegedly inconsistent rules risks degrading performance in neighbouring regions and can introduce new gaps elsewhere. In contrast, inserting targeted rules that encode the expected behaviour strengthens the rule base only where support is lacking, while preserving existing, empirically validated structures. Logical gaps are particularly problematic in systems where semantic interpretability is a core requirement.

To illustrate this phenomenon, consider a simplified FRBS designed to predict an individual's expenditure on luxury items, denoted as *AmountSpentOnLuxuryItems*, based on several inputs including *income*. Empirical evidence and intuition suggest a reasonably high positive correlation between income and expenditure on luxury goods.

Suppose the variable *income* is fuzzified into three linguistic categories: *Low*, *Medium*, and *High*. Similarly, the output variable is also divided into *Low*, *Medium*, and *High* expenditure levels. Let the rule base, constructed via an optimisation algorithm, contain only the following two rules associated with income:

Rule 1: IF income is Low THEN AmountSpentOnLuxuryItems is Low,

Rule 2: IF income is Medium THEN AmountSpentOnLuxuryItems is Medium.

Assume an input vector $\mathbf{x}$ such that $x_i = \text{income} = 6000$. Suppose that this value results in non-zero membership grades in both the *Medium* and *High* fuzzy sets, but most strongly activates the *Medium* set, thus firing Rule 2 and producing an output $o = \text{Medium expenditure}$.

Now consider increasing the income to $x_i^* = 8000$. Assume that this value lies entirely within the *High* fuzzy set. Since the rule base contains no rule for the condition *income is High*, no rule is fired for this input, and the resulting output $o^*$ defaults to a baseline or otherwise uninformed value. If

$$o^* < o,$$

this result contradicts the known positive correlation between income and expenditure.

This discrepancy illustrates a logical gap: the FRBS fails to produce an output consistent with the direction of change in an input variable that is known to be positively correlated with the output. The model fails to reflect intuitive or empirical behaviour due to incomplete rule coverage.

In this toy example, identifying the missing rule (i.e., one handling the *income is High* condition) is straightforward. However, in real-world applications where FRBS often include hundreds of rules and multiple antecedents per rule, such gaps are significantly harder to detect. The complexity arises from the combinatorial explosion of rule conditions and the non-linear interactions between multiple input variables. Furthermore, logical gaps may remain hidden until specific inputs activate missing rule conditions, which may not be evident from inspection alone.

It is also important to distinguish between logical gaps and non-linear behaviours. In some cases, the input–output relationship may not be strictly linear. For instance, if an input feature has a non-monotonic effect on the output, then a reversal in direction does not necessarily indicate a gap. Therefore, confirming the presence of a logical gap typically requires prior knowledge of the monotonic relationship between inputs and outputs, often estimated using correlation coefficients or domain knowledge.

In summary, logical gaps in FRBS arise when the rule base, due to optimisation-driven constraints, fails to represent expected monotonic relationships between inputs and outputs. These gaps can lead to counter-intuitive predictions, even in systems designed for interpretability. Detecting and correcting such gaps is essential to maintaining both the semantic consistency and user trust in predictive models.

## 4.5 Repairing Logical Gaps in Single-Antecedent FRBS

Logical gaps occur when a FRBS fails to reflect an expected monotonic relation between an input variable and the output. This section gives a systematic procedure for removing such gaps when all rules contain a single antecedent. The procedure has five main stages: *input screening*, *matrix construction*, *gap localisation*, *rule insertion*, and *weight assignment*. Each stage is discussed next, followed by an illustrative example and an algorithmic summary.

### 4.5.1 Input Screening

Only inputs with a clear monotonic influence on the output are examined for gaps. Let $y$ denote the output and $x_i$ the $i$th input feature. Compute Pearson correlations.

$$r_i = \mathrm{corr}(x_i, y), \qquad i = 1, \ldots, m.$$

Choose the set

$$\mathscr{S} = \{x_i : |r_i| \geq \tau\}$$

where $\tau \in (0,1)$ is a user-defined threshold. The absolute value filters both positively and negatively correlated inputs. Automating this step ensures that only variables capable of producing an intuitive violation are considered.

### 4.5.2 Matrix Construction

Let every input variable be partitioned into $k$ fuzzy sets, labelled generically as $\{\text{Low}, \text{Medium}, \text{High}\}$ when $k = 3$. Create an $m \times k$ matrix

$$\mathbf{A} = [a_{ij}], \qquad i = 1, \ldots, m, \ j = 1, \ldots, k,$$

called the *input–output fuzzy set matrix*. Row $i$ corresponds to input $x_i$; column $j$ corresponds to its $j$th fuzzy set $I_{ij}$. Populate each entry $a_{ij}$ with the set of output fuzzy sets

$$a_{ij} = \big\{O_\ell : \text{a rule } (I_{ij} \to O_\ell) \text{ exists}\big\}.$$

If no rule exists then $a_{ij} = \varnothing$. A compact example for $m = 6$ and $k = 3$ is shown in Table. Although the table displays linguistic labels, the procedure is independent of the actual membership functions.

### 4.5.3 Gap Localisation

A logical gap appears wherever $a_{ij} = \varnothing$ for an input $x_i \in \mathscr{S}$ while at least one other entry $a_{ik} \neq \varnothing$ exists in the same row. Multiple gaps may exist in one row. The requirement that at least one other condition be present is essential - without any supporting rules in the row, the absence of a rule cannot be interpreted as a violation of expected behaviour. If no

Table 4.1: Example input–output fuzzy set matrix ($k = 3$). The symbols $O_{\mathrm{L}}$, $O_{\mathrm{M}}$, and $O_{\mathrm{H}}$ denote low, medium, and high output fuzzy sets respectively. Empty cells indicate missing rules.

| Input | $I_{\text{Low}}$ | $I_{\text{Medium}}$ | $I_{\text{High}}$ |
|---|---|---|---|
| $x_1$ | $O_{\mathrm{H}}$ | $O_{\mathrm{M}}$ | $O_{\mathrm{L}}$ |
| $x_2$ | $O_{\mathrm{L}}$ | $\varnothing$ | $O_{\mathrm{M}}$ |
| $x_3$ | $O_{\mathrm{L}}$ | $O_{\mathrm{M}}$ | $O_{\mathrm{H}}$ |
| $x_4$ | $O_{\mathrm{L}}$ | $\varnothing$ | $O_{\mathrm{H}}$ |
| $x_5$ | $O_{\mathrm{M}}$ | $O_{\mathrm{H}}$ | $\varnothing$ |
| $x_6$ | $O_{\mathrm{L}}$ | $\varnothing$ | $\varnothing$ |

supporting rules exist for any fuzzy set of a given input feature, then the absence of a rule cannot be interpreted as a violation of expected behaviour, but only as an unmodelled region of the input space or as an input dimension that does not participate in the rule base. In this case, no directional behaviour is encoded, and therefore no inconsistency can be established. In contrast, when neighbouring or alternative output categories are already represented, the missing entry indicates that the rule base encodes some responses for that input range but fails to encode others that are required to maintain a coherent directional trend. It is this asymmetry of representation that reveals a structural deficiency rather than a simple absence of knowledge.

For single antecedent rule bases, columns are ordered by the monotone relation implied by the correlation sign.

$$\text{Low} \to \text{Medium} \to \text{High} \quad (\text{if } r_i > 0), \qquad \text{High} \to \text{Medium} \to \text{Low} \quad (\text{if } r_i < 0). \tag{4.5}$$

A missing entry breaks this ordering and signals an inconsistent semantic trajectory.

### 4.5.4 Rule Insertion

Each gap is filled by adding exactly one single-antecedent rule for $x_i$ with,

$$I_{ij} \longrightarrow O_c,$$

where $O_c$ is chosen from a candidate set that preserves the ordering in (4.5). Define the candidate set

$$\mathscr{C}_{ij} = \big\{O_\ell : O_\ell \text{ does not reverse the monotone trend}\big\}.$$

For Row 4 of Table 4.1 ($x_4$), all three output fuzzy sets satisfy the ordering, hence $|\mathscr{C}_{42}| = 3$. For Row 2 ($x_2$) the correlation is positive; only $O_L$ or $O_M$ can maintain monotonicity, so $|\mathscr{C}_{2,2}| = 2$.

When multiple consecutive gaps occur (Row 6), start from the nearest populated neighbour on one side, add a rule, and proceed sequentially to the other side. This strategy avoids abrupt jumps in the output scale.

### 4.5.5 Consequent Selection

If $|\mathscr{C}_{ij}| > 1$, a single consequent must be selected. The rule quality metric discussed in (4.2) or (4.4) or any other suitable rule metric is used to rank candidate rules. In this research (4.2) was used. The chosen rule quality metric output a score between 0 and 1 and is used to rank the candidate rules, where a higher score is better. The rule with the highest score is picked.

Alternatively, the rule quality can be used in conjunction with or in isolation as a rule picking mechanism. For instance, the rule that yields the lowest validation error can be selected. This approach is computationally heavier but directly addresses predictive accuracy.

### 4.5.6 Weight Assignment

Assign a weight $w_{\text{new}}$ to the inserted rule satisfying

$$w_{\text{new}} \leq \min\{w_r : r \in \mathscr{N}(I_{ij})\},$$

where $\mathscr{N}(I_{ij})$ is the set of neighbouring rules whose antecedents are adjacent fuzzy sets of $x_i$. This conservative choice avoids the new rule dominating the inference. Optionally, evxhaustive tuning of $(O_c, w_{\text{new}})$ pairs can be performed, but the marginal benefit must justify the computational overhead involved.

### 4.5.7 Illustrative Example

Returning to the luxury-expenditure example of Section 4.4. Assume that after optimisation the rule base lacks the rule for *income=High*. Using $k = 3$ fuzzy sets per input, construct the matrix in Section 4.5.2. Row $i$ corresponds to income; the High column is empty. Because $r_{\text{income}} > 0$, candidates are ordered Low→Medium→High, and $\mathscr{C}_{\text{income,High}} = \{O_H, O_M, O_L\}$. Compute rule quality $Q$ using (4.2) for each candidate. Suppose $O_H$ maximises $Q$. Insert the rule

IF income is High THEN AmountSpentOnLuxuryItems is High,

with weight not exceeding the weights of rules fired by *income=Medium*. Optionally, validate on a hold-out set. If RMSE improves and monotonicity is restored, retain the rule; otherwise consider the next best candidate.

Algorithm 4 summarizes the single antecedent repair mechanism.

**Algorithm 4** Repair of logical gaps in single-antecedent FRBS

**Require:** Data set $\{(\mathbf{x}^{(p)}, y^{(p)})\}_{p=1}^{N}$, rule base $\mathscr{R}$, correlation threshold $\tau$, quality metric $Q$

Compute $r_i = \text{corr}(x_i, y)$ for all inputs
$\mathscr{S} \leftarrow \{x_i : |r_i| \geq \tau\}$
Construct matrix $\mathbf{A}$ and locate gaps $a_{ij} = \varnothing$ for $x_i \in \mathscr{S}$, given $\exists a_{ik} \neq \varnothing$ in row $i$
**for all** gaps $(i, j)$ **do**
    Construct candidate set $\mathscr{C}_{ij}$ respecting ordering (4.5)
    **for all** $O_c \in \mathscr{C}_{ij}$ **do**
        Evaluate $Q$ and validation RMSE
    Select best $O_c$ and insert rule $(I_{ij} \rightarrow O_c)$
    Assign weight $w_{\text{new}} \leq \min w_{\text{neighbour}}$
Return augmented rule base $\mathscr{R}'$

### 4.5.8 Discussion

The method addresses logical gaps without altering existing rules, thus preserving interpretability. Its complexity is dominated by matrix construction $O(mk)$ and candidate evaluation $O(G \cdot N)$, where $G$ is the number of gaps. For large rule bases, a parallel implementation can reduce run-time. Although the discussed procedure targets single-antecedent rules, the matrix concept extends naturally to multiple-antecedent systems, provided that antecedent combinations are decomposed into marginal structures. This will be discussed in the next section.

## 4.6 Repairing Logical Gaps in FRBS with Multiple Antecedents

For rules with multiple antecedents, the main difference with the single antecedent scenario is the identification of the gap. For example, in a rule with 3 antecedents $A$, $B$ and $C$, if we need to examine the impact of a change in inference for $A$, where the fuzzy set changes from *medium* to *high*, we can only do so for cases where $B$ and $C$ are identical to the original rule. With this condition, the remainder of this section addresses the multiple antecedent scenario. Rules are of the form (4.1). Each input $x_i$ is partitioned into the ordered fuzzy sets $\{I_{i1}, \ldots, I_{ik}\}$. Logical gaps arise when a missing combination of antecedent sets breaks an expected monotone relation between a monotone input and the output after conditioning on the other inputs. The single-antecedent repair procedure of Section 4.5 is extended as follows.

### 4.6.1 Stage 1: Input Screening

Compute the Pearson correlations $r_i = \mathrm{corr}(x_i, y)$ on the training set. Select the monotone subset

$$\mathscr{S} = \{x_i : |r_i| \geq \tau\}, \qquad \tau \in (0,1).$$

Only these inputs are inspected for gaps. This stage is the same as the case of a single antecedent.

### 4.6.2 Stage 2: Context Construction

Let $\mathbf{j} = (j_1, \ldots, j_q)$ index a rule antecedent, with $q$ being the maximum antecedent size. Fix $x_i \in \mathscr{S}$. Hold every other antecedent at one fuzzy set to create a $(q-1)$-dimensional *context*

$$\kappa = \big(j_1, \ldots, j_{i-1}, j_{i+1}, \ldots, j_q\big),$$

For each $\kappa$ form a length-$k$ vector

$$\mathbf{A}_{i,\kappa} = [a_{i,\kappa}(j_i)]_{j_i=1}^{k}, \qquad a_{i,\kappa}(j_i) = \big\{O_\ell : \text{rule } (\mathbf{j}) \to O_\ell \text{ exists}\big\}.$$

Collecting all $\mathbf{A}_{i,\kappa}$ yields a sparse order-$q$ Boolean tensor. For reasons of aiding explainability, $q$ is usually capped to a small number such as 3 or 4.

### 4.6.3 Stage 3: Gap Localisation

Order the indices inside $\mathbf{A}_{i,\kappa}$ by

$$1 \to 2 \to \cdots \to k \ (r_i > 0), \qquad k \to k-1 \to \cdots \to 1 \ (r_i < 0).$$

A logical gap is present at $j_i$ when $a_{i,\kappa}(j_i) = \varnothing$. Only contexts with at least two neighboring populated positions are processed; this gives a defined local monotone expectation.

### 4.6.4 Stage 4: Rule Insertion

For every gap $(i, \kappa, j_i)$ define

$$\mathscr{C}_{i,\kappa,j_i} = \{O_\ell : O_\ell \text{ preserves the ordering in } \mathbf{A}_{i,\kappa}\}.$$

Insert one rule

$$\text{IF } x_1 \text{ is } A_{1j_1} \text{ AND } \dots \text{ AND } x_q \text{ is } A_{qj_q} \text{ THEN } y \text{ is } B_c, \qquad B_c \in \mathscr{C}_{i,\kappa,j_i}.$$

### 4.6.5 Stage 5: Consequent Selection

As in the single antecedent approach, use a suitable rule quality metric to select the rule with the highest quality score $Q$. Choose the consequent that maximises $Q$ and/or apply the validation-set RMSE test used for single antecedents.

### 4.6.6 Stage 6: Weight Assignment

Let $\mathscr{N}(\mathbf{j})$ be the rules that differ from $\mathbf{j}$ in exactly one antecedent index. Assign

$$w_{\text{new}} \leq \min\{w_r : r \in \mathscr{N}(\mathbf{j})\},$$

limiting the perturbation to the current inference surface.

### 4.6.7 Algorithmic Summary

Algorithm 5 outlines the procedure used to fix rational gaps when the rule base has rules with multiple antecedents.

### 4.6.8 Discussion

The main difference for multiple-antecedents is in maintaining the antecedent grouping for features identified for enforcing monotonic behaviour. Many contexts are empty in practice, further lowering complexity. The method collapses to that of Section 4.5 when $q = 1$. Existing rules stay intact, so interpretability is preserved, while local monotonicity is restored.

**Algorithm 5** Repair of logical gaps with multiple antecedents

**Require:** Data $\{(\mathbf{x}^{(p)}, y^{(p)})\}_{p=1}^{N}$, rule base $\mathscr{R}$, threshold $\tau$, quality metric $Q$
1: Compute $r_i$ and form $\mathscr{S}$
2: **for all** $x_i \in \mathscr{S}$ **do**
3:     **for all** contexts $\kappa$ **do**
4:         Build $\mathbf{A}_{i,\kappa}$ and locate gaps
5: **for all** gaps $(i, \kappa, j_i)$ **do**
6:     Generate $\mathscr{C}_{i,\kappa,j_i}$
7:     Select $O_c$ via $Q$ or validation RMSE
8:     Insert rule and set $w_{\text{new}}$ with $w_{\text{new}} \leq w_{\text{neighbour}}$
9: **return** augmented rule base $\mathscr{R}'$

### 4.6.9 Worked Example

Assume that each variable is described by the linguistic labels $\{\text{Low}, \text{Medium}, \text{High}\}$. The existing two-antecedent rule base contains the rules

$$r_1: \text{ IF income is } \textit{High} \text{ AND age is } \textit{Low} \text{ THEN savings is } \textit{High},$$

$$r_2: \text{ IF income is } \textit{Low} \text{ AND age is } \textit{Low} \text{ THEN savings is } \textit{Low}.$$

Income shows a positive monotonic relationship with savings; age does not have assumed monotonicity. The repair algorithm of Section 4.6 proceeds as follows.

**Input Screening.** The correlation test identifies

$$\mathscr{S} = \{\text{income}\}.$$

**Context Construction.** Fix the non-monotone antecedent at $\kappa = (\text{age} = \text{Low})$. For this context the income axis is represented by the length-three vector

$$\begin{aligned}\mathbf{A}_{\text{income},\kappa} = \big[&\text{savings at (income=Low, age=Low)},\\ &\text{savings at (income=Medium, age=Low)},\\ &\text{savings at (income=High, age=Low)}\big]\\ = [&\text{Low}, \varnothing, \text{High}].\end{aligned} \tag{4.6}$$

**Gap Localisation.** With a positive correlation the ordering is Low $\rightarrow$ Medium $\rightarrow$ High. The empty entry in (4.6) constitutes a logical gap because it lies between two populated neighbours.

**Candidate Set.** The neighbouring consequents are Low (left) and High (right), so any consequent that does not reverse the monotone trend is admissible:

$$\mathscr{C}_{\text{income},\kappa,2} = \{\text{Medium}, \text{High}\}.$$

**Consequent Selection.** Compute the quality score using the chosen rule quality metric $Q$.

Suppose the data gives $Q(\text{Medium}) > Q(\text{High})$. The selected consequent is therefore savings $=$ Medium.

**Rule Insertion and Weight Assignment.** Insert the rule

IF income is *Medium* AND age is *Low* THEN savings is *Medium*.

with weight $w_3 \leq \min\{w_1, w_2\}$.

**Outcome.** The updated vector becomes $[\text{Low}, \text{Medium}, \text{High}]$, restoring the expected monotonic increase in savings as income rises for individuals with low age. No existing rule is altered and interpretability is preserved.

## 4.7 Experiment to Validate the Proposed Gap Repair Method

### 4.7.1 Aim

This experiment tests whether the proposed method can detect and correct logical gaps inside a type-2 fuzzy regression rule base starting from a real-world dataset. Although the study is conducted using a type-2 system, the underlying principles and repair mechanism are equally applicable to type-1 fuzzy rule-based systems. The choice of a type-2 model is made to ensure consistency with other methods examined in this work, where type-2 fuzzy sets are employed. The chosen target task is the prediction of annual income from personal credit records. Montonic behaviour and performance are measured before and after the repair.

### 4.7.2 Dataset

Lending Club dataset was chosen to experiment with and validate our approach. This is a publicly available dataset on Kaggle [183] originally intended to predict the creditworthiness of a loan application. The dataset contains multiple continuous features. The continuous variable `Annual Income` was chosen as the output. Table 4.2 lists the thirteen input features and their linear correlations with the output as observed in the dataset.

Table 4.2: Input features and correlations with the output

| Feature | Pearson $r$ |
|---|---|
| Loan Amount | 0.42 |
| Employment Length | 0.18 |
| FICO Credit Score | 0.09 |
| Debt To Income Ratio | −0.15 |
| No. Delinquencies Last 2 Years | 0.03 |
| Earliest Credit Line Opened | −0.28 |
| No. Inquiries Last 6 Months | 0.05 |
| Months Since Last Delinquency | 0.02 |
| No. Credit Lines | 0.24 |
| Total Credit Balance | 0.41 |
| Use Of Credit Line | 0.03 |
| Total Number Credit Lines | 0.36 |
| Loan Application Description | −0.01 |

### 4.7.3 Model Construction

A type-2 fuzzy regression system was trained on the data. For this experiment, each numerical variable is partitioned into five triangular fuzzy sets labelled *very low*, *low*, *medium*, *high*, and *very high*. Let $x_j$ be the $j$th input and $\mu_{jk}(x_j)$ its membership in the $k$th fuzzy set. A rule $R_i$ takes the generic form discussed in (4.1)

A process of rule base creation using a GA to select a compact rule set was used to derive the final rule base for the model.

### 4.7.4 Identification of a Logical Gap

#### Input Screening

From table 4.2, it can be seen that most of the features have a weak linear relationship with the target. Based on the correlations, and a chosen threshold of 0.4, we identified two input features - `Loan Amount` and `Total Credit Balance` as candidates for enforcing monotonic behaviour.

**Context Construction**

| | very low | low | medium | high | very high |
|---|---|---|---|---|---|
| Total Credit Balance | [(very low, 2)] | [(very low, 2)] | [(low, 3)] | [(medium, 2)] | ∅ |
| Loan Amount | [(very low, 3)] | [(low, 5)] | [(medium, 3)] | [(high, 3)] | [(very high, 3)] |

Table 4.3: Rule Coverage Matrix

After training, the rule coverage matrix was inspected. For clarity, rule weights were assigned as discrete values normalised to the range $[1,5]$, where 1 denotes a weak rule and 5 denotes a strong rule. Each non-empty cell shows the output set chosen whenever a given input set appears in a rule antecedent. The row corresponding to the feature `Total Credit Balance` reveals that no rule involves its *very high* fuzzy set. On the other hand `Loan Amount` has no gaps and all the fuzzy sets are in a monotonic ascending order. A missing cell within `Total Credit Balance` signals a logical gap because certain input states have no representative mapping in the rule base. To see why this gap causes issues - an inference is made with a sample payload $X$ where the *high* fuzzy set of `Total Credit Balance` is triggered, and then in the next inference $X^*$, the value of `Total Credit Balance` is changed while keeping everything else constant to compare the difference in predicted output. Table 4.4 shows the original feature values used in the inference. The prediction for these inputs is $\hat{y} = 41694.7$.

For $X$, the value of `Total Credit Balance` was kept at 15840, which belonged to the core region for fuzzy set *high*. The fuzzy sets for texttttTotal Credit Balance were derived based on its distribution and can be seen in Figure 4.1. Now, in $X^*$, the value of `Total Credit Balance`: 15840 is changed to 35000, which is in the core region for fuzzy set *very high* , while keeping the other inputs unchanged. Running $X^*$ through our FRBS gives us $\hat{y}^* = 34801.76$. The prediction $\hat{y}^*$ is now lower than $\hat{y}$. Essentially, what the FRBS is predicting is, that all else being equal, the `Annual Income` will be lower if the `Total Credit Balance` increases from 15840 to 35000, which is logically anomalous. This happens as a result of a logical gap in the fuzzy sets, which has caused the centroid of the

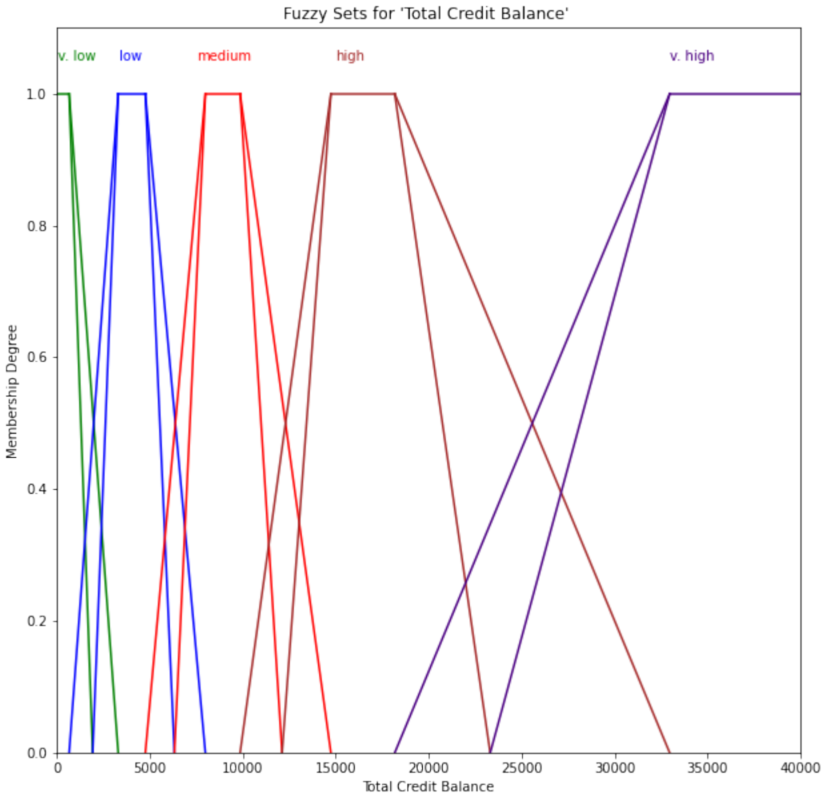


Figure 4.1: Distribution based Fuzzy Sets for Total Credit Balance

| Feature | Value |
|---|---|
| Loan Amount | 5000 |
| Employment Length | 1.0 |
| Debt To Income Ratio | 18.64 |
| No. Delinquencies (Last 2 Years) | 0 |
| Earliest Credit Line Opened | 34064.20 |
| FICO Credit Score | 700.0 |
| No. Inquiries (Last 6 Months) | 0.0 |
| Months Since Last Delinquency | 0 |
| No. Of Credit Lines | 10.0 |
| Total Credit Balance | 15840.0 |
| Use Of Credit Line | 47.1 |
| Total Credit Lines | 12.0 |
| Loan Application Description | 260 |

Table 4.4: Test Feature Values

output to shift lower as `Total Credit Balance` with its value set to 15840 fired the rule:

IF Total Credit Balance is *High* THEN Annual Income IS *Medium*

whereas `Total Credit Balance` in the domain of *very high* fuzzy set does not fire any rules.

The system outputs $\hat{y}$ and $\hat{y}^*$ display the relationship:

$$\hat{y}^* < \hat{y}$$

contradicting the observed positive correlation $r = 0.41$. The anomaly appears because *very high* fails to fire any rule, therefore other features dominate the aggregation and pull the output downwards.

### 4.7.5 Gap Repair Strategy

The proposed repair inserts at least one rule whose antecedent contains `Total Credit Balance` is one of $\{$*medium*, *high*, *very high*$\}$. A candidate consequent must respect the empirical monotonic trend of the feature with the target. Consequent sets *medium*, *high*, and *very high* all satisfy this requirement.

Each candidate rule was added with weight one and weight two to satisfy our lower or equal weight requirement. The quality score was computed using the chosen rule quality metric. The rule quality metric discussed in (4.2) was used. The effect on predictive accuracy is measured on a hold out validation set on the user's metric of choice, which in this case is RMSE. Results are summarised in Table 4.5.

Table 4.5: Test RMSE before and after repair

| Model | Weight 1 RMSE | Weight 2 RMSE |
|---|---|---|
| No change | 32 580 | 32 580 |
| *medium* added | 32 593 | 32 708 |
| *high* added | 32 427 | 32 031 |
| *very high* added | **32 233** | **31 740** |

### 4.7.6 Selection of the Final Rule

The lowest RMSE of $31,740 and the highest quality score occurs when the inserted rule is

IF Total Credit Balance is *Very High* THEN Annual Income is *Very High*,

with weight two. This rule is adopted and the model is updated with the new rule.

### 4.7.7 Validation of the Repair

The earlier counter example is re-evaluated. The prediction for $x$ remains unchanged. For $\mathbf{x}^*$, the new rule now fires for *very high* fuzzy set with $\hat{y}* = 58052.1$. For the revised output, the relationship

$$\hat{y}^* > \hat{y},$$

aligns with domain intuition and removes the previously observed logical anomaly.

### 4.7.8 Discussion

The experiment supports three claims.

1. Logical gaps can arise naturally when rule selection algorithms remove antecedent combinations.

2. A structural diagnosis based on the rule coverage matrix allows precise localisation of such gaps.

3. A simple rule insertion guided by empirical correlation reduces both the logical inconsistency and the global error metric.

It is worth noting that correcting logical gaps may lead to a deterioration in RMSE or other standard performance metrics. This occurs because the repair process introduces structural constraints that prioritise adherence to known directional relationships over strict optimisation with respect to the empirical training distribution. In regions where historical data are sparse, noisy, or biased, the original model may have achieved lower error by exploiting incidental correlations that contradict domain knowledge. Enforcing monotonic behaviour in such cases can move predictions away from locally optimal but semantically implausible fits, thereby increasing aggregate error measures.

Despite this potential cost, the trade-off is often justified in applications where interpretability, regulatory compliance, fairness, or safety are critical. Models that violate basic domain relationships may be unacceptable in operational settings even if their average predictive accuracy is marginally better. Moreover, improvements in logical consistency can enhance robustness to distribution shifts and reduce the likelihood of extreme or counter-intuitive predictions in poorly sampled regions of the input space. For these reasons, a modest increase in error metrics may be an acceptable and even desirable consequence of enforcing coherent and trustworthy model behaviour, and the final decision must reflect the priorities and risk tolerance of the deployment context.

## 4.8 Justification of the Experimental Scope

The experimental evaluation presented in this chapter is deliberately focused on a single real-world lending dataset and a targeted illustrative scenario rather than a large-scale benchmark study across multiple domains. This design choice warrants explicit justification, as it departs from the convention of broad empirical comparison that characterises much of the machine learning literature.

The primary contribution of this chapter is conceptual and methodological: the formalisation of logical gaps as a distinct category of structural deficiency in fuzzy rule-based systems, together with a principled detection and repair procedure. The experimental component serves to demonstrate the existence and practical consequence of the problem, not to claim generalised superiority of one predictive model over another. In this respect, the role of the experiment is closer to a constructive proof or worked example than to a conventional benchmark evaluation. A single, well-characterised case is sufficient to establish that logical gaps do arise under realistic conditions, that they produce semantically anomalous predictions, and that the proposed repair method can close them with measurable effect on both logical consistency and predictive accuracy.

This approach is consistent with established practice in the fuzzy systems literature, where novel inference mechanisms, rule learning strategies, and interpretability techniques are frequently introduced and validated using one or a small number of illustrative problems before broader empirical investigation is undertaken in subsequent work [171, 174, 178]. The rationale is that the explanatory value of a clearly presented single case,where every step of detection, candidate generation, selection, and validation can be traced in full outweighs the statistical coverage that a larger study would provide at the cost of reduced transparency. For a technique whose merit rests on structural reasoning about rule bases rather than on marginal improvements in aggregate error, detailed exposition of the mechanism is more informative than averaged performance tables across heterogeneous benchmarks.

Furthermore, the lending dataset was selected precisely because it embeds well-understood monotonic relationships (e.g. between total credit balance and annual income, or between loan amount and annual income) that are grounded in domain knowledge and supported by empirical correlation. This allows the logical gap and its repair to be evaluated not only numerically but also semantically, against expectations that any domain practitioner would recognise. Using a domain with clear directional priors strengthens the validity of the demonstration, since the ground truth for what constitutes a gap is unambiguous.

It is acknowledged that the limited experimental scope places certain bounds on the conclusions that can be drawn. The study does not quantify how frequently logical gaps arise across different fuzzy system configurations, nor does it assess the scalability of the repair procedure to very high-dimensional rule bases or to domains where monotonic relationships are weaker or contested. These are important questions, but they are distinct from the objective of this chapter, which is to establish the concept, formalise it, and demonstrate a viable corrective strategy. Broader empirical validation, including application to multiple datasets, alternative fuzzy system architectures, and comparison with constraint-enforcement methods embedded in the learning process, is identified as a direction for future work.

In summary, the use of a single focused experiment is a deliberate methodological choice that prioritises depth of illustration over breadth of evaluation. It enables a transparent, end-to-end demonstration of the logical gap phenomenon and its resolution, providing a concrete foundation upon which subsequent empirical studies can build.

## 4.9 Critical Analysis and Conclusion

The technique introduced here is a simple fix for a complex underlying problem. Trust in the outputs of machine learning models is a broader concern, and this research addresses only one specific aspect of that concern: logical gaps within FRBSs, applicable to both type-1 and type-2 systems.

Logical gaps occur when there is an absence of intermediate rules that would normally be expected to exist between two related rules. Such gaps can result in discontinuities in system behaviour, leading to reduced interpretability and reliability. The method proposed in this paper addresses these gaps retrospectively by identifying and inserting missing intermediate rules. While this reactive strategy mitigates specific cases of inconsistency, it does not prevent such gaps from arising in the first place.

A more comprehensive solution would involve integrating logical coherence into the rule selection process itself. This would require formalising the notion of neighbourhoods within the rule space and enforcing coverage constraints during the construction of the rule base. In this context, neighbourhoods refer to sets of rules that occupy adjacent regions in the input space and should collectively reflect a smooth and continuous mapping. Enforcing group-wise selection of such neighbouring rules could ensure logical continuity and reduce the likelihood of gaps. It is important to distinguish logical gaps from the situations typically addressed by fuzzy rule interpolation. Interpolation methods are activated when no rule sufficiently matches the current input, that is, when rule coverage is absent. In contrast, logical gaps arise in regions where rules do fire and produce outputs, but where the resulting behaviour violates known directional relationships. Since standard interpolation mechanisms are not triggered under such conditions, they cannot correct monotonicity violations caused by inconsistent or insufficient rule structures.

Implementing constraints based on the notion of rule neighbourhood introduces challenges. It increases the complexity of the rule selection process, requiring mechanisms to identify neighbourhood structures and manage trade-offs between coverage, interpretability, and rule base size. It may also necessitate changes to the underlying learning algorithm or optimisation framework used to generate the rule base.

### 4.9.1 Contribution

This research proposes a method to address logical gaps in FRBSs by retrospectively identifying missing intermediate rules and inserting them into the rule base. This approach enhances the interpretability and reliability of the system by reducing discontinuities in behaviour, offering a targeted mitigation for inconsistencies that arise from absent rules. The proposed technique supports the broader goal of responsible machine learning by enabling practitioners to deploy interpretable models that behave in a manner consistent with established domain knowledge, thus fulfilling the key objective of this research.

## Summary

This chapter examined the problem of *logical gaps* in fuzzy rule-based systems (FRBSs) for regression, defined as structural violations of known monotonic or directional relationships between inputs and outputs. Such gaps arise when optimisation-driven rule selection prioritises compactness and global error metrics, inadvertently removing intermediate or low-support rules that are nevertheless necessary for preserving coherent semantic behaviour. As a result, small changes in inputs may produce implausible output reversals or discontinuities, undermining interpretability, trust, and regulatory acceptability.

To address this issue, the chapter proposed a targeted, retrospective repair method that detects and corrects logical gaps without modifying existing validated rules. The approach screens for monotonic input variables, constructs structured representations of rule coverage, localises missing rule configurations, and inserts minimal corrective rules that preserve expected directional trends. The method is formulated for both single-antecedent and multiple-antecedent rule bases through the use of context-conditioned neighbourhood analysis, ensuring that monotonic behaviour is enforced locally while maintaining overall model transparency.

An experimental study on a real-world lending dataset demonstrated that a small number of carefully selected rule insertions can eliminate substantial logical inconsistencies, often with little or no degradation in predictive performance and in some cases with measurable improvement. The chapter concludes that interpretability alone is insufficient if internal logical relationships are not respected, and that lightweight structural repair provides a practical path toward more trustworthy and semantically consistent fuzzy regression models, while highlighting future directions for integrating neighbourhood-aware constraints directly into rule learning.

# Chapter 5

# Hybrid Fuzzy Regression with Enhanced Accuracy and High Explainability

## 5.1 Introduction

Despite the advantages of fuzzy regression approaches, balancing interpretability and accuracy remains challenging. This research addresses this gap by proposing a novel framework: the Hybrid IT2 Mamdani–TSK Fuzzy System (HIT2-MTSK), which integrates the strengths of both Mamdani and TSK paradigms.

The proposed approach contributes as follows:

- Integrates the linguistic interpretability of Mamdani systems with the numerical precision of TSK models.
- Introduces a new rule type containing both a fuzzy component and a functional component that produces a crisp output value derived from antecedent features rather than from the centroid of the consequent fuzzy set.
- Employs two dominance types for rule evaluation, one associated with the fuzzy consequent and another with the crisp consequent.

This framework aims to accurately handle complex real-world datasets while preserving interpretability and adaptability, directly addressing regulatory requirements for explainability [184]. The research advances fuzzy regression modelling and contributes to broader machine learning by providing a robust, interpretable predictive framework.

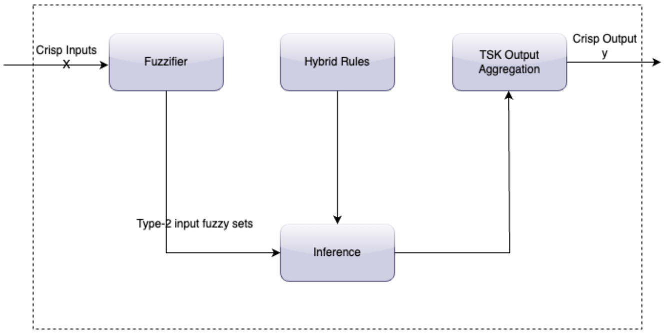


Figure 5.1: Structure of Hybrid Mamdani-TSK System

The HIT2-MTSK methodology was evaluated using benchmark datasets. Across six datasets, the proposed method achieved the best fuzzy-system performance in four cases and the best overall performance in one case, with RMSE improvements ranging from 0.4% to 19%, while maintaining full interpretability.

In Section 5.2, the HIT2-MTSK framework is formally presented. Section 5.3 reports experimental evaluations, followed by limitations, contributions, and future research directions in Section 5.5.

## 5.2 The Proposed Hybrid Mamdani-TSK System

The proposed system - Hybrid IT2 Mamdani TSK fuzzy System (HIT2-MTSK), integrates IT2 FRBSs and TSK approaches to address regression challenges by combining interpretability and precision. HIT2-MTSK introduces a novel rule structure that maintains linguistic interpretability whilst achieving enhanced numerical accuracy through constrained polynomial functions.

The steps involved in the HIT2-MTSK flow are shown in Figure 5.1:

- Fuzzification of inputs and target with IT2 Fuzzy Sets
- Rule Generation with Mamdani and TSK Components
- Rule Base selection using Ant Colony Optimization (ACO)
- Defuzzification

This section uses a running example with two sample features (cement, blast furnace slag) from the Concrete dataset to illustrate the relevant concepts. The Concrete dataset

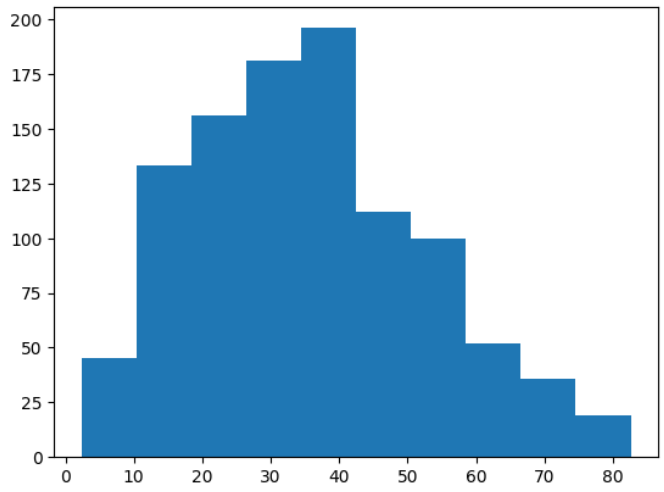


Figure 5.2: Distribution for Concrete Compressive Strength

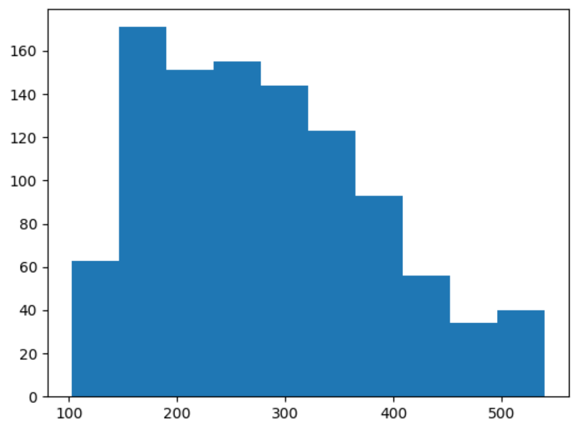


Figure 5.3: Distribution for Concrete Cement

serves as one of the test datasets for evaluating the HIT2-MTSK method, with detailed information provided in Section 5.3. The distributions of the regression target (compressive strength) and the selected features for this dataset are as shown in Figure 5.2, 5.3 and 5.4.

Three fuzzy sets - Low, Medium and High, were used to fuzzify the variables as shown in Figure 5.6.

### 5.2.1 Fuzzification with Interval Type-2 Fuzzy Sets

Inputs and outputs are fuzzified using IT2 fuzzy sets, which better represent uncertainty by allowing membership values to vary within an interval. The mathematical formulation is presented in Equation 2.4, which is discussed in Section 2.2.2. IT2 sets handle uncertainty

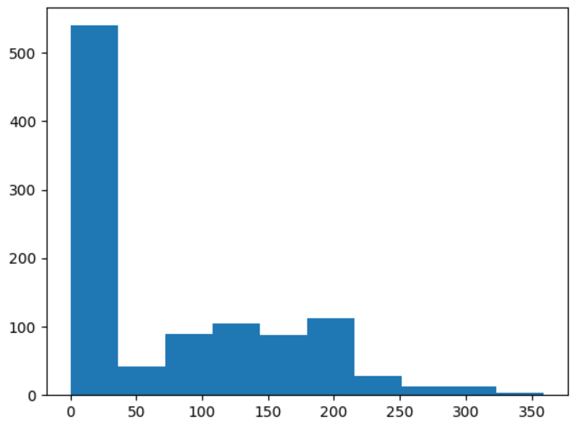


Figure 5.4: Distribution for Blast Furnace Slag

more flexibly than type-1 fuzzy sets, enhancing robustness for real-world applications [72].

The fuzzification process follows a data-driven approach to partition the input and output spaces. For each variable $v_i$ with domain $[v_{i,min}, v_{i,max}]$, the algorithm determines partition points based on data distribution characteristics. The three-partition approach of using 'Low', 'Medium' and 'High' fuzzy sets ensures that the divisions align with human understanding and provide meaningful linguistic descriptions of the data [63]. However, more partitions can be used if deemed appropriate for the use case.

The fuzzy sets are generated based on the data distribution. The extreme left and extreme right fuzzy sets are modelled as IT2 left shoulder and IT2 right shoulder respectively. All other fuzzy sets are modelled as IT2 trapezoidal fuzzy sets as shown in Figure 5.6. To ensure that the linguistic terms Low, Medium, and High reflect meaningful data characteristics, the parameters for each fuzzy set are determined based on the underlying variables' distribution.

For Concrete Compressive Strength in our example, the lower and upper bound for each fuzzy set (where the support is greater than zero) are, Low: [2.33, 32.04], Medium: [16.89, 54.9] and High: [37.91, 82.6] as per the points highlighted in Figure 5.5.

### 5.2.2 Rule Generation with Hybrid Components and Rule Dominance

The rule generation process represents the core innovation of this research. Unlike traditional approaches that generate either Mamdani or TSK rules, the HIT2-MTSK method creates hybrid rules that combine the interpretability of linguistic consequents with the precision

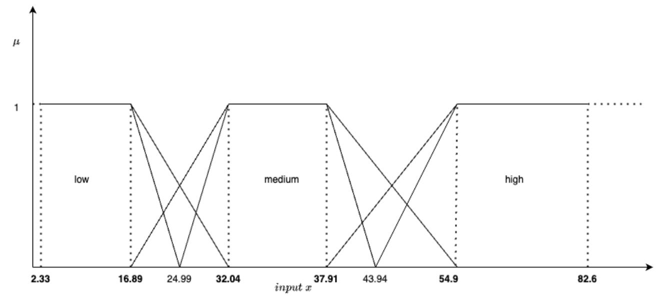


Figure 5.5: Significant Fuzzy Set Points for Concrete Strength

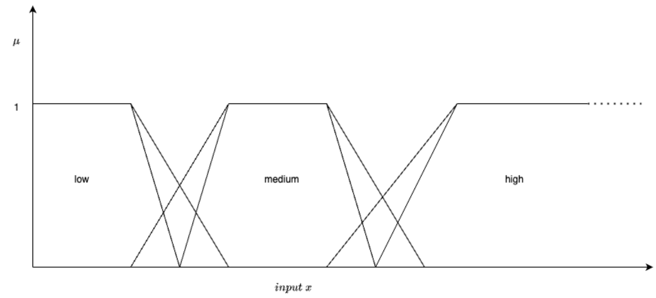


Figure 5.6: Schema for Interval Type-2 Fuzzy Sets

of mathematical functions. This novel hybridization allows the system to maintain human-readable explanations while achieving the computational accuracy typically associated with TSK models. By integrating both paradigms within individual rules, the approach overcomes the conventional trade-off between interpretability and performance.

**Hybrid Rule Structure**

Each rule $R_j$ in the HIT2-MTSK system follows the structure:

$$R_j : \text{IF } x_1 \text{ is } \tilde{F}_{j1} \text{ AND ... AND } x_n \text{ is } \tilde{F}_{jn} \text{ THEN } (G_j, f_j(\mathbf{x})),$$

where $\tilde{F}_{ji}$ are IT2 fuzzy sets, $G_j$ is the linguistic consequent (IT2 fuzzy set), and $f_j(\mathbf{x})$ is the TSK function component.

Each rule in the HIT2-MTSK hybrid methodology has two components for the consequent:

1. **Fuzzy Component:** This is the same as the consequent in a Mamdani FRBS – i.e. each rule points to an output fuzzy set, as shown in Eq. (2.1). This component preserves the linguistic interpretability of the rule.

2. **TSK Function Component:** The second component is a function that maps antecedent features to target values. A crisp output is provided using a n-order polynomial equation. However, there are two notable departures from the traditional TSK methodology:

   - The equation is only trained on a subset of training samples (rows) which have positive firing-strength for the fuzzy component, and
   - the output from this equation is forced to be bound within the lower and upper bounds of the associated fuzzy set.

**TSK Function Component Design**

The function component supports both linear functions $n = 1$ and higher-degree polynomials $n > 1$ to handle non-linear mappings. Higher-degree polynomials can be used without sacrificing explainability because the rules maintain consistency with their fuzzy consequents throughout the mapping process. The general form of the function for degree $n$ is:

$$f_j(\mathbf{x}) = w_0 + \sum_{i=1}^{m} w_i x_i + \sum_{i=1}^{m}\sum_{k=i}^{m} w_{ik} x_i x_k + \ldots + \sum_{i_1, i_2, \ldots, i_n} w_{i_1 i_2 \ldots i_n} \prod_{l=1}^{n} x_{i_l},$$

where $w$ represents the coefficient terms and $m$ is the number of input variables.

For instance, for inputs $x_1$ and $x_2$, and n set to 2, the function component takes the form:

$$y = w_0 + w_1 x_1 + w_2 x_2 + w_3 x_1^2 + w_4 x_2^2 + w_5 x_1 x_2,$$

where $w_0, w_1, \ldots, w_5$ are coefficients learned from the data. The TSK function component allows the model to capture complex non-linear relationships and makes the prediction system potentially much more accurate when compared to other fuzzy approaches as will be discussed in the results section.

**Constraint Mechanism**

To ensure that the meaning of the Mamdani consequent is maintained, we need to make sure that the output from the TSK component is always within the bounds of the consequent fuzzy set. The constraint mechanism operates as follows:

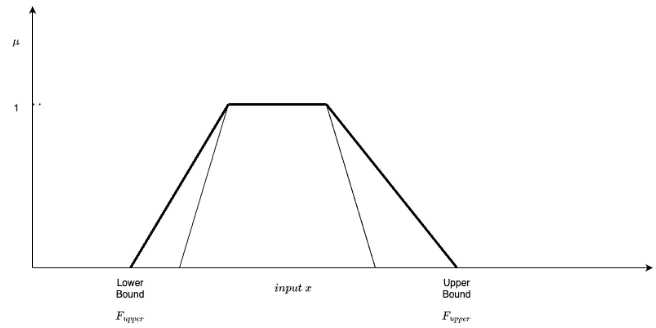


Figure 5.7: Lower and Upper Bounds of a Trapezoid Type-2 Fuzzy set

$$y_{\text{constrained}} = \begin{cases} \text{LowerBound}(F_{upper}), & \text{if } f_j(\mathbf{x}) < \text{LowerBound}(F_{upper}) \\ f_j(\mathbf{x}), & \text{if LowerBound}(F_{upper}) \leq f_j(\mathbf{x}) \leq \text{UpperBound}(F_{upper}) \\ \text{UpperBound}(F_{upper}), & \text{if } f_j(\mathbf{x}) > \text{UpperBound}(F_{upper}) \end{cases}$$

where $F_{upper}$ is the upper type-1 fuzzy set for the underlying IT2 fuzzy set as shown in Figure 5.7. This maintains interpretability for the underlying rule by keeping the output within the bounds of the underlying fuzzy set and prevents unreasonable extrapolations.

**Rule Generation Algorithm**

**Algorithm 6** Hybrid Rule Generation

**Input:** Training data $D = \{(\mathbf{x}_i, y_i)\}_{i=1}^{N}$, IT2 fuzzy sets
**Output:** Rule base $\mathscr{R} = \{R_1, R_2, \ldots, R_K\}$
Generate all possible rule antecedent combinations
**for** each antecedent combination $A_j$ **do**
  Calculate firing strength for all data points
  Identify subset $D_j \subset D$ where firing strength $> 0$
  Determine consequent fuzzy set $G_j$ based on $D_j$
  Train TSK function $f_j$ using only $D_j$:

$$\mathbf{w}_j = \arg\min_{\mathbf{w}} \sum_{(\mathbf{x},y) \in D_j} (f_j(\mathbf{x}; \mathbf{w}) - y)^2$$

  Apply constraint mechanism to ensure outputs within $G_j$ bounds
  Calculate rule quality metrics (support, confidence, RMSE)
  **if** rule meets quality threshold **then**
    Add rule $R_j = (A_j, G_j, f_j)$ to rule base $\mathscr{R}$

The complete rule generation process follows the steps shown in 5.2.2. For the concrete dataset example, for variables *concrete_cement* and *concrete_blast_furnace_slag,* a sample rule generated in our models is:

$$
\begin{aligned}
&\text{IF cement is High AND blast\_furnace\_slag is High} \\
&\text{THEN } \big(\text{compressive\_strength\_High}, \\
&\quad y = 0.3 \cdot \text{cement} - 0.6 \cdot \text{blast\_furnace\_slag} \\
&\quad - 5.29 \times 10^{-4} \cdot (\text{cement})^2 \\
&\quad + 1.68 \times 10^{-3} \cdot \text{cement} \cdot \text{blast\_furnace\_slag} \\
&\quad + 2.82 \times 10^{-4} \cdot (\text{blast\_furnace\_slag})^2\big)
\end{aligned}
\tag{5.1}
$$

The consequent of rule shown in Eq. (5.1) is a 2-tuple. The first part of the tuple is the fuzzy set that the rule points to (*High* in the example shown), the second part of the tuple is the TSK function equation which is used to derive the output within the bounds of the fuzzy set to which the rule points. In this example, the output from the equation will always be bound within the lower and upper bounds of the Fuzzy set 'High' of the Concrete Compressive Strength.

As an instance, if Cement value is 375 and Blast Furnace Slag is 300 – both values are in the 'High' fuzzy set of their features which means that our sample rule will be triggered – the output from our sample rule, based on the equation shown in Eq. (5.1) is 72.62, which is inside the lower and upper bound of the High fuzzy set for Compressive Strength and hence does not need changing. Assuming that there was a case in which the rule output was 85 - i.e. outside the upper bound of 82.6, in this case the rule output would be changed to 82.6 and similarly if there was a case in which the rule output was lower than 37.91, then the rule output will be changed to 37.91.

**Computational Complexity (Maximum of Three Antecedents)**

Let $n$ denote the total number of input variables, $k$ the number of fuzzy sets per variable, and $N$ the number of training samples. Each rule is restricted to at most three antecedents, meaning that only subsets of input variables of size $s \in \{1,2,3\}$ are considered.

For a fixed subset size $s$, the number of possible antecedent combinations is given by selecting $s$ variables from $n$ and assigning one of $k$ fuzzy sets to each selected variable, yielding

$$R_s = \binom{n}{s} k^s.$$

Thus, the total number of candidate rules is

$$R_{\leq 3} = \sum_{s=1}^{3} \binom{n}{s} k^s.$$

For each candidate rule, the following operations are performed:

- **Firing strength computation:** For each of the $N$ samples, the membership values of $s \leq 3$ antecedents are combined, giving a cost of $O(Ns) = O(N)$.
- **Subset selection and consequent estimation:** Identifying the subset $D_j$ and estimating the consequent fuzzy set $G_j$ both require linear scans of the data, resulting in $O(N)$.
- **TSK consequent training:** For rules with $s$ antecedents and TSK polynomial degree $d$, the number of regression parameters is

  $$m_s = \binom{s+d}{d}.$$

  Estimation via least-squares regression using QR decomposition requires

  $$O(|D_j| \cdot m_s^2) \leq O(N \cdot m_s^2),$$

  which dominates the per-rule computational cost.
- **Constraint enforcement and quality metrics:** Applying output constraints and computing support, confidence, and RMSE are linear in $|D_j|$ and bounded by $O(N)$.

Hence, the total worst-case computational complexity of the rule generation procedure is

$$O\left(N \sum_{s=1}^{3} \binom{n}{s} k^s m_s^2\right), \quad \text{where } m_s = \binom{s+d}{d}.$$

Since $s$ is bounded by 3, this yields a polynomial complexity with respect to $n$:

$$O\left(N n^3 k^3 \binom{3+d}{d}^2\right),$$

which is significantly lower than the exponential complexity $O(k^n)$ obtained when all $n$ input variables are used in each rule.

In practice, pruning based on minimum support and quality thresholds further reduces the effective number of trained rules, leading to substantially lower runtime than the worst-case bound.

### Numerical Example (Maximum of Three Antecedents)

Consider a dataset with $n = 10$ input variables, $k = 3$ fuzzy sets per variable, and $N = 1000$ training samples. Restricting each rule to at most three antecedents, the total number of candidate rules is

$$R_{\leq 3} = \binom{10}{1}3^1 + \binom{10}{2}3^2 + \binom{10}{3}3^3 = 30 + 405 + 3240 = 3675.$$

**Quadratic TSK consequent ($d = 2$).** For rules with at most three antecedents, the maximum number of regression parameters occurs when $s = 3$:

$$m_3 = \binom{3+2}{2} = \binom{5}{2} = 10.$$

The dominant regression cost per rule is therefore

$$O(Nm_3^2) \approx 1000 \cdot 10^2 = 1.0 \times 10^5,$$

**Cubic TSK consequent ($d = 3$).** The number of parameters for $s = 3$ becomes

$$m_3 = \binom{3+3}{3} = \binom{6}{3} = 20.$$

The dominant regression cost per rule is

$$O(Nm_3^2) \approx 1000 \cdot 20^2 = 4.0 \times 10^5,$$

This example shows that limiting the number of antecedents per rule converts the exponential growth in rule count into polynomial growth and significantly reduces the regression cost by limiting the number of input variables in each TSK consequent.

### 5.2.3 Rule Weights

In a novel departure from other fuzzy regression techniques, this research calculates two weights for each rule - both derived from the training data. The dual-weight system enables the model to leverage both the linguistic characteristics of fuzzy rules and their numerical performance. The use of linguistically motivated rule weights in this work is well aligned with existing research on weighted fuzzy rule-based systems. In particular, the fuzzy weighting mechanisms employed here share conceptual foundations with the fuzzy weighted adaptive boosting framework [185], where rule weights reflect both representational reliability and data-driven relevance. Similarly, the integration of rule weights based on fuzzy support and confidence is consistent with the weighted inference paradigm proposed in W-Infer-Polation [186], which emphasises the role of weights in balancing interpretability and approximation quality during reasoning. Earlier works [187] further supports this distinction by highlighting the transition from purely approximative fuzzy classifiers toward more descriptive models, where linguistic fidelity and coverage are explicitly quantified. Collectively, these studies provide a strong theoretical basis for the proposed dual-weight scheme, in which fuzzy dominance captures linguistic adequacy while the error-based dominance independently encodes numerical predictive performance.

The first technique is based on [188] and calculates the rule weight based on the fuzzy support and confidence of the rule. The second weight is based on the accuracy of the TSK rule as calculated on the training data.

**Fuzzy Rule Weight Calculation**

The fuzzy weight captures the linguistic relevance and coverage of a rule within the dataset. This method calculates fuzzy dominance for rules by combining support and confidence values derived from the fuzzy membership intervals of antecedents and consequents. The process involves several steps:

- **Calculate Firing Strength (FS):** The function starts by determining the firing strength of the rule's antecedent for each data instance. Each premise in the antecedent contributes to the firing strength using a chosen t-norm (e.g., minimum or product).

  For a given data instance $x$ and a rule with an antecedent having $n$ premises, where each premise $A_i$ is a Type-2 fuzzy set with membership grade $[\mu_{A_{i1}}(x), \mu_{A_{i2}}(x)]$, the firing strength interval $[f_1(x), f_2(x)]$ is calculated using a t-norm $T$:

$$[f_1(x), f_2(x)] = T([\mu_{A_{11}}(x), \mu_{A_{12}}(x)], ..., [\mu_{A_{n1}}(x), \mu_{A_{n2}}(x)])$$

For the product t-norm:

$$f_1(x) = \prod_{i=1}^{n} \mu_{A_{i1}}(x), \quad f_2(x) = \prod_{i=1}^{n} \mu_{A_{i2}}(x).$$

- **Support of Antecedent and Rule:** The support of the antecedent represents the average membership degree across the dataset:

$$\text{Support}(A_j) = \frac{1}{N} \sum_{p=1}^{N} \mu_{A_j}(x_p).$$

For support of the Rule: The firing strength is combined with the consequent's membership interval values to compute the rule's support. The support and confidence equations 5.2 and 5.3 relate to the $j^{th}$ rule with Antecedent $A_j$ and consequent fuzzy set $\tilde{C}_j$ for a dataset with $N$ instances with $x_p = (x_{p1}, ..., x_{pn})$ and $p = 1, 2, ..., N$,

$$\text{Support}(A_j \rightarrow \tilde{C}_j) = \frac{1}{|N|} \sum_{p=1}^{N} \mu_{A_j}(x_p) \cdot \mu_{C_j}(y_p). \tag{5.2}$$

For IT2 fuzzy sets, the interval support is calculated as:

$$[S_{lower}, S_{upper}] = \left[ \frac{1}{N} \sum_{p=1}^{N} f_1(x_p) \cdot \underline{\mu}_{C_j}(y_p), \frac{1}{N} \sum_{p=1}^{N} f_2(x_p) \cdot \bar{\mu}_{C_j}(y_p) \right].$$

where $f_1(x_p)$ and $f_2(x_p)$ represent the lower and upper firing strengths of the antecedent respectively.

- **Fuzzy Confidence:** The rule's confidence is the ratio of its support to the antecedent's support.

$$\text{Confidence}(A_j \rightarrow \tilde{C}_j) = \frac{\sum_{p=1}^{N} \mu_{A_j}(x_p) \cdot \mu_{C_j}(y_p)}{\sum_{p=1}^{N} \mu_{A_j}(x_p)}. \tag{5.3}$$

For IT2 sets, the confidence interval is:

$$[C_{lower}, C_{upper}] = \left[ \frac{\sum_{p=1}^{N} f_1(x_p) \cdot \underline{\mu}_{C_j}(y_p)}{\sum_{p=1}^{N} f_2(x_p)}, \frac{\sum_{p=1}^{N} f_2(x_p) \cdot \bar{\mu}_{C_j}(y_p)}{\sum_{p=1}^{N} f_1(x_p)} \right].$$

- **Dominance Calculation:** The fuzzy weight of the rule is calculated taking into account both the fuzzy support and fuzzy confidence of the rule. For example, the dominance can be calculated as:

$$D_{fuzzy} = [S_{Rule\_lower} \cdot C_{lower}, S_{Rule\_upper} \cdot C_{upper}].$$

This ensures the fuzzy weight reflects both how well the rule fits the data (confidence) and its prevalence (support). The main role of fuzzy rule weights is at the time of generation of a viable universe of rules.

**Error-based Dominance**

The error-based dominance provides a performance-oriented weight that directly reflects the rule's prediction accuracy.

- At the time of rule generation, each rule is evaluated using its performance on training data, quantified as its RMSE over applicable instances, wherever the firing strength of the rule is non-zero. The RMSE for rule $j$ is calculated as:

$$RMSE_j = \sqrt{\frac{1}{|D_j|} \sum_{(\mathbf{x},y)\in D_j} (f_j(\mathbf{x}) - y)^2},$$

  where $D_j$ is the subset of training data with positive firing strength for rule $j$.

  The error-based dominance of a rule is then calculated as:

$$\text{Dominance}_{error}^{j} = \frac{1}{1 + RMSE_j}, \tag{5.4}$$

where $RMSE_j$ is the RMSE of the j$^{\text{th}}$ rule. Rules with lower RMSE have higher dominance values, allowing them to contribute more significantly during the rule selection process and at time of inference. This ensures the dynamic prioritization of accurate rules.

### Integration of Dual Weights

The dual-weight system provides two distinct measures of rule quality. The fuzzy weight captures the linguistic relevance and coverage of a rule within the dataset, whilst the error-based dominance directly reflects the rule's prediction accuracy. Both weights are calculated during rule generation and stored with each rule.

The fuzzy rule weights are used at the time of generation of the universe of rules, pointing to a fuzzy consequent set. Whereas, error-based dominance allows rules to be ranked and selected during the rule selection process with ACO.

In our example, the rule shown in Eq. (5.1) has two weights, the fuzzy weight is [0.436, 0.457] and the TSK component dominance is 0.065 – which is calculated using its RMSE in Eq. (5.4). The RMSE for this rule is 14.3 and the error-based dominance can be calculated from this as

$$\text{Dominance}_{error}^{sample} = \frac{1}{1+14.3} = 0.065.$$

## 5.2.4 Subset Selection Using Ant Colony Optimization

After rule generation, a subset of rules is selected using ACO. The selection process addresses the challenge of finding an optimal balance between rule base compactness and prediction accuracy. ACO optimizes the trade-off between model compactness and accuracy by iteratively refining a subset of rules. The fitness function evaluates rule subsets based on the combined RMSE of selected rules on training data with an optimization goal of minimizing the RMSE on the training/validation dataset. The ACO algorithm is inspired by the foraging behaviour of ants, who use pheromones to mark paths leading to food sources.

### ACO Implementation

The optimization procedure starts with an initial set of rules, and each ant in the colony constructs a solution by selecting a subset of these rules. The specific implementation follows these steps:

The ants evaluate the performance (RMSE) of their solutions on a training dataset. Based on the performance, they deposit pheromones on the rules they selected. Rules that lead to better solutions receive more pheromones, making them more likely to be selected in future

**Algorithm 7** ACO for Rule Selection

**Input:** Rule base $\mathscr{R}$, parameters $\{\alpha, \beta, \rho, Q, num\_ants, max\_iter\}$
**Output:** Selected rule subset $\mathscr{R}^*$
Initialize pheromone matrix $\tau_{ij} = \tau_0$
Initialize best solution $\mathscr{R}^* = \emptyset$, best_RMSE $= \infty$
**for** iteration = 1 to $max\_iter$ **do**
    **for** ant = 1 to $num\_ants$ **do**
        Construct solution $\mathscr{R}_{ant}$ by:
            Start with empty rule set
            Select rules probabilistically:
            $P_j = \frac{[\tau_j]^\alpha \cdot [\eta_j]^\beta}{\sum_{k \in \text{unselected}} [\tau_k]^\alpha \cdot [\eta_k]^\beta}$
            where $\eta_j = D^j_{\text{error}}$
        Evaluate $RMSE(\mathscr{R}_{ant})$
        Update local best if improved
    Update pheromones:
        $\tau_j = (1 - \rho) \cdot \tau_j$
        $\tau_j = \tau_j + \sum_{ant} \Delta\tau_j^{ant}$ if $j \in \mathscr{R}_{ant}$
        $\Delta\tau_j^{ant} = Q / RMSE_{ant}$
    Update global best if improved

iterations. The pheromone levels on the rules evaporate over time, and the process repeats for a fixed number of iterations. The algorithm keeps track of the best solution (set of rules) found so far and its corresponding RMSE. The full process is shown in Figure 5.8.

### Solution Construction

Each ant constructs a solution by:

1. Starting with an empty rule set
2. Iteratively adding rules based on selection probabilities that balance:
    - Historical success (pheromone values $\tau$)
    - Individual rule accuracy (heuristic information $\eta$)
3. Stopping when a size constraint is met or no improvement occurs

### Computational Complexity

The ACO algorithm has a complexity of $O(max\_iter \cdot num\_ants \cdot |\mathscr{R}|^2 \cdot N_{val})$ where:

- $max\_iter$ is the maximum number of iterations

- *num_ants* is the number of ants
- $|\mathscr{R}|$ is the size of the initial rule base
- $N_{val}$ is the size of the validation set

ACO is computationally efficient, making it well-suited for high-dimensional datasets [189], while ensuring the selected rule subset retains high predictive performance. The ACO implementation in this research follows the following broad steps.

- The algorithm runs multiple iterations till a maximum number of iterations is reached or the score stops improving for a pre-defined number of iterations.
- Each ant constructs a solution (a subset of rules) probabilistically using selection probabilities for each rule by balancing two factors: (1) historical success (represented by "pheromone" values) and (2) individual rule accuracy (prioritising rules with lower errors). The system then picks a group of rules within a predefined size range, favouring those with higher probabilities.
- The solution is evaluated using RMSE on training and validation data.
- Identifies the best solution (lowest RMSE) in the current iteration.
- Updates the global best solution if an improvement is found.
- If no improvement occurs for a set number of iterations (patience), the loop breaks early.
- Adjusts pheromone levels based on solution quality.
- Log the current iteration and best RMSE.
- After completion, the best-performing rules are extracted from the best solution.

**Computational Complexity (ACO for Rule Selection)**

Let $K = |\mathscr{R}|$ be the number of candidate rules, $N$ the number of data samples, $A = num\_ants$ the number of ants per iteration, $T = max\_iter$ the number of iterations, and $S$ the average number of rules selected by each ant.

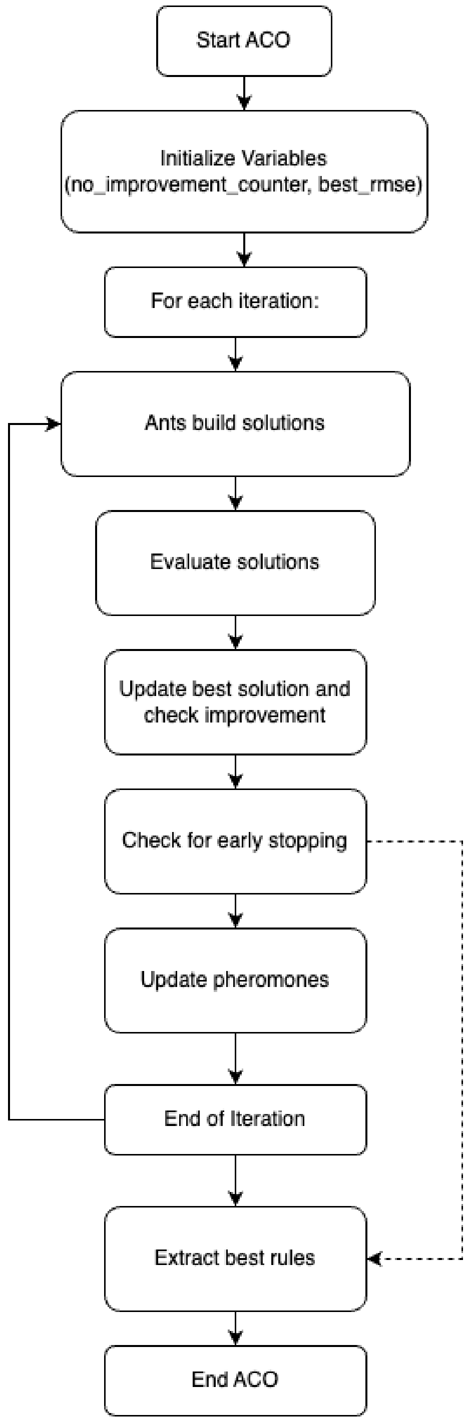


Figure 5.8: Ant Colony Optimisation Overview

In each iteration, every ant performs two main operations: (i) constructing a subset of rules using pheromone-guided probabilistic selection, and (ii) evaluating the selected subset by computing the prediction error (RMSE) over the dataset.

**Solution construction.** Selecting $S$ rules from $K$ candidates requires computing and sampling from probability distributions over the remaining rules. Depending on the implementation, this step typically costs between $O(K+S)$ and $O(K+S\log K)$ per ant, and in a naive implementation up to $O(SK)$. Since this step does not depend on the dataset size, it is generally not the dominant cost when $N$ is large.

**RMSE evaluation.** To evaluate a candidate solution, predictions must be computed for all $N$ samples using the selected $S$ rules. If evaluating one rule on one sample has constant cost, the evaluation step costs $O(NS)$ per ant. If each rule evaluation involves checking a small number of antecedent conditions, this cost becomes $O(NSM)$, where $M$ is the number of antecedents per rule, typically a small constant.

**Pheromone update.** Pheromone evaporation is applied to all $K$ rules, and deposition is applied to the rules selected by the ants, resulting in a cost of $O(K+AS)$ per iteration, which is negligible compared to solution evaluation when $N$ is large.

**Overall complexity.** Combining these steps, the total computational complexity is dominated by the RMSE evaluation and can be expressed as

$$O(T \cdot A \cdot N \cdot S \cdot M),$$

where $M$ is small and often treated as a constant. In practical settings where $N \gg K$, the evaluation of candidate solutions dominates the total runtime.

**Numerical Example (ACO for Rule Selection)**

Consider a rule base with $K = 500$ candidate rules and a dataset of $N = 10{,}000$ samples. Assume $A = 30$ ants are used per iteration for $T = 50$ iterations, and that each ant selects on average $S = 50$ rules per solution. Assume each rule evaluation involves checking $M = 3$ antecedent conditions.

**RMSE evaluation cost.** For a single ant, evaluating one solution requires

$$N \cdot S \cdot M = 10{,}000 \cdot 50 \cdot 3 = 1.5 \times 10^6$$

basic rule evaluations. Across all ants and iterations, this becomes

$$T \cdot A \cdot N \cdot S \cdot M = 50 \cdot 30 \cdot 10{,}000 \cdot 50 \cdot 3 = 2.25 \times 10^9.$$

**Solution construction and pheromone updates.** Selecting rules probabilistically typically requires at most a few thousand operations per ant, and pheromone updates require only a few thousand operations per iteration. Even when aggregated over all ants and iterations, these costs are on the order of $10^6$ operations, which is negligible compared to the billions of operations required for RMSE evaluation.

This example illustrates that, in practical settings with large datasets, the overall runtime of the ACO-based rule selection procedure is dominated by repeated evaluation of candidate rule subsets, and scales approximately linearly with the number of samples, ants, iterations, and selected rules.

### 5.2.5 Defuzzification Using TSK Weighted Mean

The defuzzification process combines outputs from multiple fired rules to produce a final crisp prediction. Unlike traditional Mamdani systems that require complex defuzzification methods, HIT2-MTSK leverages the TSK framework's computational efficiency whilst maintaining interpretability.

#### Inference Process

For a given input vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, the inference process to predict the output $\hat{y}$ consists of the following steps:

1. **Rule Activation:** Each rule antecedent is defined by a conjunction of interval type-2 (IT2) fuzzy sets,

$$R_i : \text{If } x_1 \text{ is } \tilde{F}_{i1} \text{ and } \ldots \text{ and } x_n \text{ is } \tilde{F}_{in}, \text{ then } y = f_i(\mathbf{x}).$$

For each input variable $x_j$, its degree of membership in the IT2 fuzzy set $\tilde{F}_{ij}$ is given by an interval

$$\mu_{\tilde{F}_{ij}}(x_j) = [\mu_{ij}^L(x_j), \mu_{ij}^U(x_j)],$$

which reflects uncertainty in the membership function.

The firing strength of rule $R_i$ is obtained by aggregating the antecedent memberships using a t-norm operator $T(\cdot)$ (product or minimum in this work):

$$[\omega_i^L, \omega_i^U] = T\big(\mu_{\tilde{F}_{i1}}(x_1), \mu_{\tilde{F}_{i2}}(x_2), \ldots, \mu_{\tilde{F}_{in}}(x_n)\big),$$

where the t-norm is applied to the lower and upper membership bounds separately, producing an interval-valued firing strength that captures the uncertainty in rule activation.

2. **Type Reduction:** To enable numerical aggregation in the TSK framework, the interval firing strength must be converted into a crisp value. Instead of performing iterative type-reduction procedures such as the Karnik–Mendel (KM) algorithm, the mean of the lower and upper bounds is used:

$$\omega_i = \frac{\omega_i^L + \omega_i^U}{2}.$$

This approximation is computationally efficient and yields a stable estimate of rule activation, particularly when the footprint of uncertainty (FOU) is symmetric or when lower and upper membership functions are derived from bounded parameter perturbations. Such mean-based approximations to type-reduction have been explored in the literature as efficient alternatives to full iterative methods. Non-iterative techniques such as the Nie–Tan approach and related closed-form type-reduction methods compute representative outputs by averaging upper and lower bounds of the output IT2 set, yielding results close to those of the more costly Karnik–Mendel (KM) algorithms while significantly reducing computation time [190].

3. **Rule Output and Constraint Mechanism:** Each rule produces a crisp output through its TSK consequent function $f_i(\mathbf{x})$. However, since the rule is also associated with a consequent fuzzy set $G_i$ (derived from training samples covered by the rule), the output

is constrained to remain within the support of the upper membership function of $G_i$ in order to preserve semantic consistency between antecedent and consequent.

The constrained output is computed as:

$$y_{\text{constrained}} = \begin{cases} \text{LowerBound}(F_{upper}), & \text{if } f_i(\mathbf{x}) < \text{LowerBound}(F_{upper}) \\ f_i(\mathbf{x}), & \text{if LowerBound}(F_{upper}) \leq f_i(\mathbf{x}) \leq \text{UpperBound}(F_{upper}) \\ \text{UpperBound}(F_{upper}), & \text{if } f_i(\mathbf{x}) > \text{UpperBound}(F_{upper}) \end{cases}$$

where $\text{LowerBound}(F_{upper})$ and $\text{UpperBound}(F_{upper})$ denote the lower and upper limits of the support of the upper membership function of the IT2 consequent fuzzy set $G_i$.

This mechanism prevents extrapolated outputs that are inconsistent with the data region represented by the rule and improves robustness by limiting extreme predictions, especially when polynomial consequents are used.

4. **Weighted Aggregation:** All fired rules ($\omega_i > 0$) contribute to the final prediction through a weighted average, consistent with the TSK inference scheme:

$$\hat{y}_{\text{final}} = \frac{\sum_{i=1}^{M} \omega_i \, y_{\text{constrained},i}}{\sum_{i=1}^{M} \omega_i},$$

where $M$ is the number of activated rules for the given input, and $\omega_i$ is the type-reduced firing strength of rule $R_i$.

This inference procedure preserves the uncertainty modeling capability of IT2 fuzzy sets during rule activation while maintaining efficient and numerically stable prediction through simplified type reduction and constrained TSK consequents.

### Handling Rule Conflicts

When multiple rules fire with similar activation strengths but different consequents, the weighted averaging naturally resolves conflicts. The contribution of each rule is proportional to its firing strength, ensuring smooth transitions between different regions of the input space.

**Computational Efficiency**

For a single input instance, the overall inference cost in the proposed TSK-based fuzzy system is dominated by rule evaluation and aggregation. Computing firing strengths requires evaluating membership functions and applying a t-norm across $n$ antecedents for each of the $M$ fired rules, resulting in a cost of $O(M \cdot n)$. Defuzzification is then performed using a weighted average of rule outputs, which has linear cost $O(M)$ and is negligible compared to rule evaluation.

In contrast, Mamdani-type fuzzy systems require aggregation of output membership functions followed by centroid defuzzification, which typically involves numerical integration or discretization over the output domain. The computational cost therefore depends on the resolution of the output universe and is substantially higher than the simple weighted averaging used in TSK inference.

**Example Calculation (IT2–TSK Inference with Constraints)**

Consider an input instance with

$$x_{\text{cement}} = 400, \qquad x_{\text{slag}} = 250.$$

Assume two IT2–TSK rules are activated from the selected rule subset $\mathscr{R}^*$:

- **Rule 1:** *High cement* AND *High slag* $\Rightarrow y = f_1(\mathbf{x})$, with consequent IT2 set $G_1$.
- **Rule 2:** *High cement* AND *Medium slag* $\Rightarrow y = f_2(\mathbf{x})$, with consequent IT2 set $G_2$.

**1) IT2 membership evaluation.** For each antecedent, the interval type-2 membership is represented by a lower and upper membership function (LMF/UMF). For the given input, suppose the membership intervals are:

$$\begin{aligned}
\mu_{\widetilde{HighCement}}(400) &= [0.75,\ 0.90], \\
\mu_{\widetilde{HighSlag}}(250) &= [0.60,\ 0.80], \\
\mu_{\widetilde{MedSlag}}(250) &= [0.20,\ 0.40].
\end{aligned}$$

**2) Interval firing strengths.** Using a product t-norm (applied separately to LMF and UMF), the interval firing strengths are

$$[\omega_1^L, \omega_1^U] = \Big[0.75 \cdot 0.60,\ 0.90 \cdot 0.80\Big] = [0.45,\ 0.72],$$
$$[\omega_2^L, \omega_2^U] = \Big[0.75 \cdot 0.20,\ 0.90 \cdot 0.40\Big] = [0.15,\ 0.36].$$

**3) Type reduction of firing strengths (mean).** The interval firing strengths are converted to crisp activations using the mean:

$$\omega_1 = \frac{\omega_1^L + \omega_1^U}{2} = \frac{0.45 + 0.72}{2} = 0.585,$$
$$\omega_2 = \frac{\omega_2^L + \omega_2^U}{2} = \frac{0.15 + 0.36}{2} = 0.255.$$

**4) TSK consequent outputs.** Evaluate each rule's TSK consequent function (e.g., polynomial) at the input:

$$f_1(\mathbf{x}) = 75.3, \qquad f_2(\mathbf{x}) = 68.5.$$

**5) Constraint mechanism.** Each rule is associated with a consequent IT2 fuzzy set $G_i$. Let the support bounds of the *upper* membership function of $G_i$ be:

$$[\text{LowerBound}(F_{upper,1}), \text{UpperBound}(F_{upper,1})] = [70,\ 80],$$

$$[\text{LowerBound}(F_{upper,2}), \text{UpperBound}(F_{upper,2})] = [60,\ 72].$$

Using the constraint mechanism:

$$y_{\text{constrained}} = \begin{cases} \text{LowerBound}(F_{upper}), & \text{if } f_i(\mathbf{x}) < \text{LowerBound}(F_{upper}) \\ f_i(\mathbf{x}), & \text{if LowerBound}(F_{upper}) \le f_i(\mathbf{x}) \le \text{UpperBound}(F_{upper}) \\ \text{UpperBound}(F_{upper}), & \text{if } f_i(\mathbf{x}) > \text{UpperBound}(F_{upper}) \end{cases}$$

we obtain:

$$y_{\text{constrained},1} = 75.3 \quad (\text{already within } [70, 80]), \qquad y_{\text{constrained},2} = 68.5 \quad (\text{already within } [60, 72]).$$

**6) Weighted aggregation (final prediction).** The final TSK prediction is computed via weighted averaging:

$$y_{\text{final}} = \frac{\omega_1 y_{\text{constrained},1} + \omega_2 y_{\text{constrained},2}}{\omega_1 + \omega_2} = \frac{0.585 \cdot 75.3 + 0.255 \cdot 68.5}{0.585 + 0.255}.$$

Compute the numerator and denominator:

$$0.585 \cdot 75.3 = 44.0505, \qquad 0.255 \cdot 68.5 = 17.4675, \qquad 0.585 + 0.255 = 0.84,$$

thus

$$y_{\text{final}} = \frac{44.0505 + 17.4675}{0.84} = \frac{61.518}{0.84} \approx 73.24.$$

This example demonstrates the full IT2–TSK inference flow: interval-valued membership grades produce interval firing strengths, which are type-reduced to crisp activations. The TSK consequents provide numerical precision, while the constraint mechanism ensures semantic consistency with the consequent fuzzy sets, and the weighted aggregation yields the final prediction.

## 5.3 Experiments and Results

This section evaluates the performance of the HIT2-MTSK fuzzy system for regression analysis. The experiments assess the effectiveness of combining Mamdani interpretability with TSK precision across diverse benchmark datasets. The evaluation demonstrates how the HIT2-MTSK approach balances accuracy and interpretability compared to existing fuzzy and non-fuzzy methods.

### 5.3.1 Dataset Description

Six benchmark datasets from the KEEL repository were utilised to evaluate the HIT2-MTSK approach [191]. These datasets represent diverse application domains and vary in complexity, providing comprehensive evaluation scenarios.

**Concrete Compressive Strength Dataset**

This dataset contains 1,030 instances with 8 input variables representing concrete mixture components including cement, water, and aggregates. The target variable is compressive strength, measured in megapascals.

**Diabetes Dataset**

The Diabetes dataset comprises 768 records with 2 input attributes. The target variable represents a quantitative measure of diabetes progression one year after baseline. This medical dataset presents challenges in capturing physiological interactions within a regression framework.

**ELE-2 Dataset**

With 11,105 instances and 4 predictor variables, this dataset addresses energy load forecasting. The large sample size and temporal nature of the data provide insights into the scalability of the proposed approach.

**Mortgage Dataset**

This financial dataset contains 1,045 instances with 14 variables for predicting loan delinquency risk. The dataset includes borrower characteristics and loan details, representing a high-dimensional regression problem in the financial domain.

**Treasury Dataset**

Comprising 956 records with 16 financial attributes derived from time-series data, this dataset targets yield rate prediction. The high dimensionality and financial context test the method's capability in economic modelling applications.

**Wankara Dataset**

This meteorological dataset contains 1,609 records of weather information from Ankara spanning from 01/01/1994 to 28/05/1998. The regression task involves predicting daily mean temperature based on 4 input variables.

### 5.3.2 Experimental Setup

The experiments followed a standardised evaluation protocol using 5-fold cross-validation as provided by the KEEL repository. This ensured a fair comparison with benchmark results.

Two variants of the HIT2-MTSK model were evaluated:

- **HIT2-MTSK-D2**: HIT2-MTSK with degree 2 polynomial functions
- **HIT2-MTSK-D3**: HIT2-MTSK with degree 3 polynomial functions

The HIT2-MTSK methods were compared against five benchmark algorithms [191,192]:

- **MP**: Multilayer perceptron (non-fuzzy)
- **SMOreg**: Sequential Minimal Optimisation for regression (non-fuzzy)
- **WM**: Wang and Mendel's fuzzy algorithm
- **CHV**: Cordón, Herrera and Villar's fuzzy algorithm
- **GLD-WM**: Granularity with global lateral parameters and rule base by Wang-Mendel

Root Mean Square Error (RMSE) served as the primary evaluation metric, calculated as the average across the 5-fold cross-validation. The ACO algorithm parameters were set with 50 ants, 100 iterations maximum, pheromone influence $\alpha = 1$, heuristic influence $\beta = 2$, and evaporation rate $\rho = 0.1$.

### 5.3.3 Experimental Setup

The experiments followed a standardized evaluation protocol using 5-fold cross-validation as provided by the KEEL repository. This protocol is widely adopted in the fuzzy systems and machine learning literature and ensures direct comparability with published benchmark results. The dataset is randomly partitioned into five equally sized folds; in each run, four folds are used for training and one fold is used for testing, and the final performance is reported as the average over the five test folds, thereby reducing variance due to data partitioning.

Two variants of the proposed HIT2-MTSK model were evaluated in order to analyze the impact of consequent model complexity:

- **HIT2-MTSK-D2**: HIT2-MTSK with quadratic (degree 2) TSK polynomial consequents, offering a balance between model flexibility and computational efficiency.

- **HIT2-MTSK-D3**: HIT2-MTSK with cubic (degree 3) TSK polynomial consequents, providing higher representational capacity at the cost of increased computational complexity.

The proposed methods were compared against five representative benchmark algorithms commonly used in regression and fuzzy modeling tasks [191, 192]:

- **MP**: Multilayer Perceptron, representing non-fuzzy neural regression models.
- **SMOreg**: Sequential Minimal Optimization for regression, representing kernel-based learning methods.
- **WM**: Wang and Mendel's fuzzy rule-based learning algorithm, a classical data-driven fuzzy modeling approach.
- **CHV**: Cordón, Herrera and Villar's fuzzy rule generation algorithm, which applies genetic learning to optimize fuzzy partitions and rule bases.
- **GLD-WM**: A granularity-based fuzzy modeling method that combines global lateral tuning with Wang–Mendel rule extraction.

**Evaluation Metric.** Root Mean Square Error (RMSE) was used as the primary evaluation metric and was computed as the average over the five cross-validation folds. RMSE is a standard performance measure for regression tasks because it directly reflects the magnitude of prediction errors in the same physical units as the target variable, and it penalizes large deviations more strongly due to the squared error term. This property is particularly important in engineering datasets, such as concrete strength prediction, where large errors are more critical than small fluctuations. Moreover, RMSE is the default metric reported in the KEEL repository for regression benchmarks, ensuring fair and consistent comparison with previously published results.

**ACO Parameter Settings.** The Ant Colony Optimization (ACO) algorithm was employed for rule subset selection, and its parameters were chosen following commonly recommended practices in the ACO and combinatorial optimization literature [193, 194] as well as prior applications of ACO to subset selection in machine learning [195].

The number of ants was set to 50, representing a moderate colony size that provides sufficient population diversity for exploring alternative rule subsets while maintaining reasonable computational cost. The maximum number of iterations was set to 100 to allow pheromone trails to stabilize without excessive runtime.

The pheromone influence parameter was set to $\alpha = 1$, assigning moderate importance to accumulated search experience, while the heuristic influence parameter was set to $\beta = 2$, giving stronger guidance to rule quality heuristics during solution construction, a common configuration in practical ACO implementations [194]. The pheromone evaporation rate was set to $\rho = 0.1$, which helps prevent premature convergence by reducing excessive reinforcement of early solutions while still promoting high-quality rule subsets [196].

Overall, these parameter values fall within commonly reported ranges for ACO-based combinatorial optimization and provided stable convergence behavior across all evaluated datasets.

### 5.3.4 Results

Table 5.1 presents the RMSE values for all methods across the six datasets. The HIT2-MTSK variants demonstrate competitive performance, achieving the best fuzzy method results in 4 out of 6 datasets.

Table 5.1: RMSE comparison across benchmark datasets

| Method | Concrete | Diabetes | ELE-2 | Mortgage | Treasury | Wankara |
|---|---|---|---|---|---|---|
| MP | 7.86 | 0.63 | 158.05 | 0.11 | 0.25 | 1.61 |
| SMOreg | 8.45 | 0.65 | 210.87 | 0.15 | 0.55 | 1.58 |
| WM | 21.21 | 0.69 | 545.38 | 0.92 | 0.74 | 2.66 |
| CHV | 10.13 | 0.68 | 356.18 | 0.21 | 0.42 | 2.26 |
| GLD-WM | 7.32 | 0.70 | 245.76 | 0.17 | 0.35 | 1.95 |
| HIT2-MTSK-D2 | 7.61 | 0.79 | 196.48 | 0.15 | 0.35 | **1.58** |
| HIT2-MTSK-D3 | **7.29** | 0.80 | **189.28** | **0.13** | **0.27** | **1.58** |

### 5.3.5 Discussion

The experimental results reveal several key findings regarding the HIT2-MTSK approach. On the Concrete dataset, HIT2-MTSK-D3 achieves the overall best performance with an RMSE of 7.29, outperforming both fuzzy methods (GLD-WM: 7.32) and non-fuzzy approaches (MP: 7.86). This demonstrates the effectiveness of combining interpretable fuzzy rules with precise TSK functions.

For the Diabetes dataset, both HIT2-MTSK variants show higher RMSE values (HIT2-MTSK-D2: 0.79, HIT2-MTSK-D3: 0.80) compared to non-fuzzy methods. The limited number of input features (2) in this dataset may contribute to this performance gap, suggesting the approach contributes more in higher-dimensional problems.

The ELE-2 dataset results show significant improvement over traditional fuzzy methods. HIT2-MTSK-D3 achieves an RMSE of 189.28, representing a 65% improvement over WM and 47% over CHV. While non-fuzzy MP performs better (158.05), the HIT2-MTSK method substantially narrows the accuracy-interpretability gap.

Financial datasets demonstrate strong performance for the hybrid approach. On the Mortgage dataset, HIT2-MTSK-D3 produces the best fuzzy method result (0.13), closely approaching the non-fuzzy MP performance (0.11). The Treasury dataset shows similar patterns, with HIT2-MTSK-D3 (0.27) outperforming all fuzzy baselines and approaching MP (0.25).

The Wankara dataset yields identical performance for both HIT2-MTSK variants (1.58), matching SMOreg and surpassing all fuzzy baselines. This suggests that increased polynomial complexity may not always help in improving performance for data with inherent variability.

The degree-3 variant (HIT2-MTSK-D3) generally outperforms degree-2 (HIT2-MTSK-D2), except for the Diabetes dataset where D2 shows marginally better results. This indicates that higher-order polynomial functions capture complex relationships effectively but may risk over-fitting on smaller or simpler datasets.

Overall, the HIT2-MTSK approach achieves the best fuzzy method performance in 66.7% of datasets (4 out of 6), outperforms non-fuzzy methods in 2 datasets, and provides the overall best result for the Concrete dataset. The RMSE improvements range from 0.4% to 19% compared to the best existing fuzzy methods, demonstrating substantial gains while maintaining rule interpretability.

The success of the hybrid approach stems from its ability to maintain Mamdani-style linguistic rules whilst incorporating TSK precision through bounded polynomial functions. This design preserves the explainability essential for fuzzy systems whilst achieving competitive accuracy against opaque models.

## 5.4 Case Study: Predicting California Housing Prices

The effectiveness of the proposed method can be demonstrated through its application to the California Housing dataset, a widely used benchmark in machine learning and predictive modelling research. This dataset, originally derived from the 1990 U.S. Census, contains information about housing prices and their relationship to demographic and geographic features across various districts in California. The proposed HIT2-MTSK method will be compared against benchmark results including results from a Mamdani IT2 regression model - demonstrating the value that HIT2-MTSK adds over the Mamdani model. The task consists of predicting the median house value in each district based on explanatory variables such as median income, average number of rooms, average household size, population, and geographical coordinates. This prediction problem is not only of academic interest but also of practical relevance, as accurate housing price estimation supports decision-making in areas such as urban planning, real estate investment, and policy design. The California Housing dataset is particularly suitable for evaluating the proposed method for two main reasons. First, it is a real-world dataset with inherent complexity, including non-linear relationships, heterogeneous feature distributions, and geographic variability. Second, the dataset has been extensively studied in the literature, which allows for a meaningful comparison with baseline models and previously reported results. In the following subsections, we describe the dataset characteristics, experimental setup, results, and discussion of findings, including a comparison with the Mamdani approach applied to the same dataset, ultimately demonstrating the practical applicability of the proposed method in a real-world prediction task.

### 5.4.1 Dataset Description

The California Housing dataset originates from the 1990 U.S. Census and has become a standard benchmark for regression tasks in machine learning. It contains 20 640 observations, each representing a California district, and is described by 8 predictive features along with one target variable. The features include both demographic and geographic attributes, which are summarized as follows:

- **MedInc**: median income in the district (in tens of thousands of US dollars).
- **HouseAge**: median age of houses in the district.

- **AveRooms**: average number of rooms per household.
- **AveBedrms**: average number of bedrooms per household.
- **Population**: total population of the district.
- **AveOccup**: average household size (occupants per household).
- **Latitude**: geographical latitude of the district.
- **Longitude**: geographical longitude of the district.

The target variable is the **median house value** in each district, expressed in units of 100 000 US dollars. Observations were randomly partitioned into training and testing subsets, with 80% of the data used for training and 20% reserved for evaluation. This partitioning ensures that the model is evaluated on unseen samples, providing a reliable measure of generalization performance.

### 5.4.2 Results and Discussion

Table 5.2 presents the predictive performance of the HIT2-MTSK method in comparison with the baseline models on the California Housing dataset. Results are reported for Root Mean Squared Error(RMSE). The results demonstrate that the HIT2-MTSK method outperforms the other explainable baseline models on RMSE, including Mamdani FRBS. Compared to traditional linear regression, the reduction in error highlights the ability of the method to capture non-linear relationships in the data. Opaque approaches deliver a lower RMSE, however lack explainability at the level of training data as well as individual inferences.

Table 5.2: Comparison of predictive performance on the California Housing dataset.

| **Model** | **RMSE** |
|---|---|
| Linear Regression (LR) | 0.728 |
| CART (RF) | 0.720 |
| NAM | 0.562 |
| EBM | 0.557 |
| XGBoost | 0.532 |
| DNN | 0.492 |
| Mamdani FRBS | 0.751 |
| HIT2-MTSK | 0.695 |

### 5.4.3 Evaluation Metrics for Explainability

Drawing from the comprehensive evaluation framework proposed by Vilone and Longo [197], we selected metrics applicable to direct fuzzy regression systems. Whilst their framework targets post-hoc wrapper models that explain black-box systems, several of their metrics - particularly those assessing rule complexity, coverage, and robustness - translate directly to evaluating the inherent interpretability of our fuzzy ruleset. The HIT2-MTSK methodology was examined using five key metrics:

- **Classes Covered** measures whether rules exist for different output ranges. Based on the prediction range in the test data (0.78 to 4.93), the model covers 96% of the actual test data range (0.15 to 5.00), with slight compression at the extremes. The model achieved 100% coverage in terms of the output fuzzy sets, with rules defined for Low, Medium, and High house value ranges.

- **Active Rules per Prediction** counts rules with firing strength exceeding various thresholds. At a threshold of 0.15, an average of 8.38 rules fire per prediction. This drops to 6.33 at 0.25 and 3.83 at 0.5. The decrease at higher thresholds shows that whilst multiple rules contribute to each prediction, only a few dominate with high firing strength.

- **Rule Base Characteristics** comprise two measures. The rule base contains 75 rules total, each with an average of 2.67 antecedents. These values indicate a compact, interpretable model where individual rules remain simple enough for human comprehension.

- **Dataset Coverage** confirms that 100% of test instances activate at least one rule. This complete coverage ensures the model can handle all input combinations without gaps.

- **Robustness to Noise** was assessed by adding Gaussian noise at three levels to input features. The noise was added proportionally based on the standard deviations of each feature in the training dataset. The mean absolute change in predictions, expressed as a percentage of the mean house value, was 1.18% at 1% noise, 5.84% at 5% noise, and 12.24% at 10% noise levels. The near-linear relationship between noise level and prediction change demonstrates stable behaviour without sudden discontinuities.

These metrics collectively demonstrate that HIT2-MTSK maintains the interpretability expected of fuzzy systems whilst achieving comprehensive coverage and robust predictions. The gradual decrease in active rules at higher firing thresholds shows that whilst multiple rules contribute to each prediction, a smaller subset dominates the decision-making process.

### 5.4.4 Comparison with Mamdani FRBS

A primary goal of this research was to surpass Mamdani FRBS performance while preserving its level of explainability. The HIT2-MTSK models achieved their results using a compact rule-base containing 75 rules. Although Mamdani FRBS models were trained with the same number of rules, they struggled to achieve comparable predictive accuracy. To ensure fair comparison, both models used identical configurations: the same number of input and output fuzzy-sets and the same maximum rule length. Figure 5.9 displays the relationship between predicted and actual house values for the testing dataset using the HIT2-MTSK approach. Figure 5.10 shows the corresponding scatter plot for the same data using the FRBS model. These figures clearly demonstrate that the hybrid approach produces a much tighter clustering around the ideal prediction line. The Mamdani FRBS scatter plot reveals distinct horizontal banding patterns at positions corresponding to the output fuzzy set centroids, visible as horizontal structures in Figure 5.10. This banding occurs in Mamdani systems when all (or most) rules for a given prediction activate the same output fuzzy-set, causing predictions to cluster within a narrow range. The hybrid approach eliminates this effect by generating predictions across the entire support region of the output fuzzy-sets, resulting in a more uniform distribution of predicted values.

The residual distribution for the HIT2-MTSK approach (Figure 5.11) shows errors distributed approximately symmetrically around zero, indicating no significant systematic bias. The Mamdani approach's residual distribution (Figure 5.12) similarly shows no bias, but exhibits a wider error spread, confirming the superior performance of the hybrid HIT2-MTSK approach.

The experimental results confirm the proposed method's effectiveness in predicting housing prices and demonstrate clear improvements over the traditional Mamdani approach. Its superior performance compared to established baselines confirms both its robustness and practical applicability for real-world regression tasks.

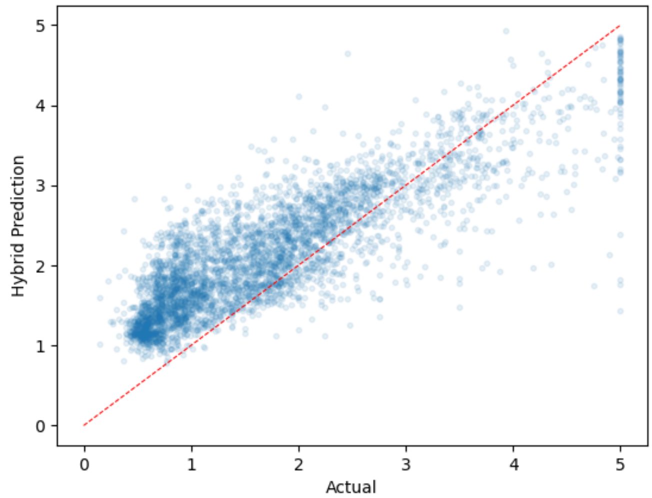


Figure 5.9: Scatter plot of actual versus predicted median house values using HIT2-MTSK

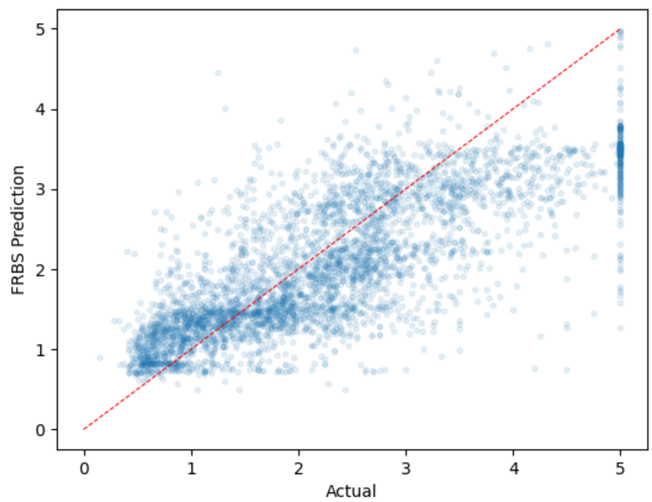


Figure 5.10: Scatter plot of actual versus predicted median house values using a Mamdani FRBS

## 5.5 Contributions, Limitations and Further Research

### 5.5.1 Research Contributions

#### Methodological Innovations

HIT2-MTSK presents three primary methodological contributions to fuzzy regression modelling. First, it introduces a novel hybrid rule structure that combines Mamdani interpretabil-

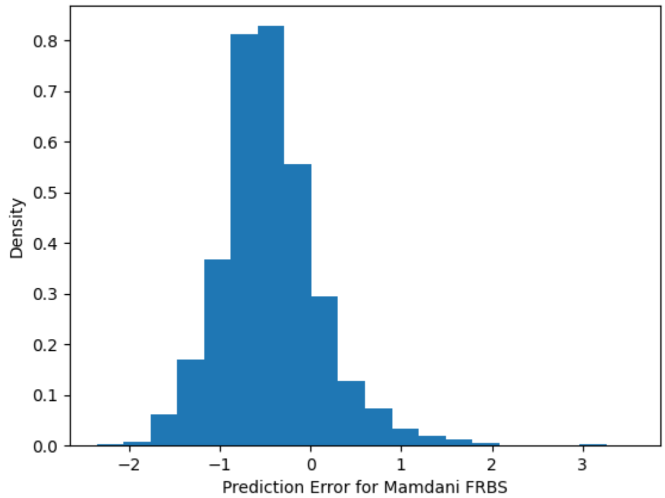


Figure 5.11: Residual distribution for HIT2-MTSK

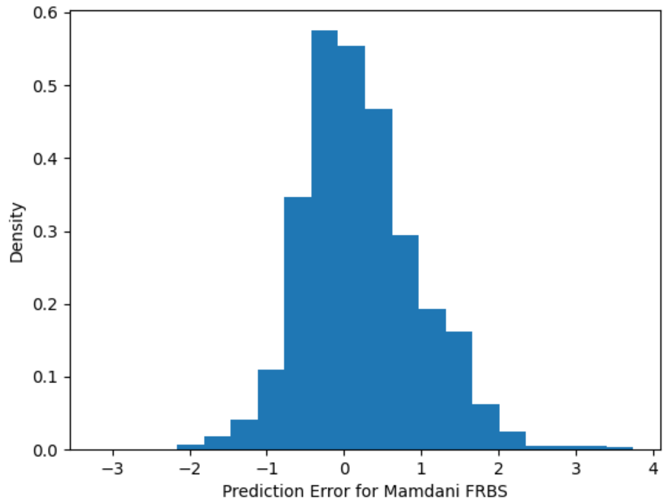


Figure 5.12: Residual distribution for Mamdani FRBS

ity with TSK precision through dual-component consequents. Each rule maintains a fuzzy component for linguistic interpretation whilst incorporating a polynomial function for precise output calculation. Unlike traditional approaches that force a choice between interpretability and accuracy, this method preserves linguistic clarity whilst improving predictive performance.

Second, the incorporation of dual dominance weights enables dynamic rule prioritisation. The fuzzy weight, calculated using support and confidence measures (5.2 and 5.3), captures rule prevalence in the data. The error-based dominance (5.4) prioritises rules based on prediction accuracy. This dual-weighting scheme allows the system to balance rule interpretability with performance during inference.

Third, the constraint mechanism ensures TSK outputs remain within fuzzy set boundaries. This preserves semantic consistency between rule components and prevents unrealistic extrapolations that could compromise interpretability.

#### Addressing the Interpretability-Accuracy Gap

The HIT2-MTSK approach addresses a significant gap in fuzzy system literature. Traditional Mamdani systems suffer from limited precision due to centroid-based defuzzification, whilst TSK systems sacrifice interpretability for accuracy [2]. Previous attempts to hybridise these approaches often resulted in complex architectures that compromised both objectives [66].

This research demonstrates that these limitations are not inherent trade-offs. By maintaining the linguistic structure of Mamdani rules whilst enhancing output precision through constrained TSK functions, the method achieves competitive performance without sacrificing explainability. The results show RMSE improvements ranging from 0.4% to 19% across benchmark datasets.

### 5.5.2 Critical Analysis

#### Performance Assessment

The experimental results reveal both strengths and limitations of the HIT2-MTSK approach. The method achieves best-in-class fuzzy performance in 4 out of 6 tested datasets (66.67%) and outperforms all benchmarks in the Concrete dataset. Performance analysis reveals domain-specific variations. Financial datasets (Mortgage, Treasury) show strong results, with the hybrid approach closely matching non-fuzzy methods. Engineering datasets (Concrete, ELE-2) demonstrate the method's ability to capture complex non-linear relationships.

#### Algorithmic Design Choices

The choice of ACO for rule selection was driven by its relative efficiency over other evolutionary approaches. Whilst ACO provides efficient subset selection, its stochastic nature may lead to suboptimal solutions in complex search spaces. The algorithm's sensitivity to parameter settings ($\alpha$, $\beta$, $\rho$) introduces additional hyperparameter tuning requirements. The method's performance on low-dimensional datasets (Diabetes: 2 features) suggests potential overfitting when feature spaces are restricted. The polynomial functions may lack sufficient input diversity to generalise effectively. This suggests that optimal polynomial complexity is dataset-dependent, with simpler datasets performing better with a less complex approach. However, this also means that a higher complexity polynomial can not be taken as granted as the better choice and thus is a candidate for optimisation.

### 5.5.3 Limitations and Future Directions

#### Computational Complexity and Scalability

The training processes of evolutionary FRBSs are computationally intensive offline due to the requirement of iterative computation. In order to obtain the (nearly) optimal solution, evolutionary FRBS require all the training data to be presented, and it can take a much longer time for evolutionary algorithms to converge depending on the nature of the problem and the scale of training data. Although not specific to our research, this is a general constraint impacting evolutionary FRBSs. For high-dimensional datasets, this becomes prohibitive.

#### Structural Limitations

The choice of fixed three-partition fuzzy structure (Low, Medium, High) and the use of trapezoidal membership functions as fixed choices have worked well for these use-cases but are not necessarily the most optimal choices. Gaussian or asymmetric membership functions might prove to be better candidates for capturing certain data characteristics.

#### Future Research Directions

There are several areas in which further research could benefit the HIT2-MTSK approach. First, adaptive fuzzy partitioning methods could adjust granularity based on local data density and feature importance. Information-theoretic measures could guide partition boundary placement. Similarly, experiments with the choice of fuzzy sets would potentially improve the results.

Second, alternative optimisation techniques may improve efficiency. Gradient-based methods could optimise polynomial coefficients directly, whilst evolutionary strategies might better handle the discrete rule selection problem. Hybrid optimisation combining local and global search could balance exploration and exploitation. The choice of polynomial complexity can also be exposed as an optimisation parameter.

Third, the framework should be extended to support explicit uncertainty quantification. Prediction intervals could be derived by exploiting both the interval-valued firing strengths of IT2 fuzzy rules and the dispersion of rule-level predictions. For example, lower and upper predictions could be computed by propagating the lower and upper firing strengths through the weighted TSK aggregation, yielding interval-valued outputs that reflect antecedent uncertainty. Furthermore, uncertainty arising from polynomial extrapolation beyond well-supported data regions could be modeled by incorporating distance-based confidence measures or variance estimates from local regression residuals. Such interval or probabilistic predictions would provide users with reliability indicators and support decision-making in risk-sensitive applications.

Fourth, multi-objective optimisation frameworks should explicitly balance interpretability and accuracy. Metrics quantifying rule base coherence, redundancy, and coverage would support this approach.

Finally, the method should be extended to handle missing data and temporal dependencies. Online learning capabilities would enable adaptation to changing data distributions whilst maintaining interpretable rule structures.

## Chapter Summary

This chapter presented the Hybrid Interval Type-2 Mamdani–TSK Fuzzy System (HIT2-MTSK), a novel regression framework designed to achieve a balanced integration of interpretability and predictive accuracy. By combining linguistic Mamdani-style fuzzy consequents with numerical TSK polynomial consequents within a unified hybrid rule structure, the proposed approach preserves human-readable rule semantics while enabling precise numerical modeling. The use of interval type-2 fuzzy sets further enhances robustness by explicitly modeling uncertainty in membership functions, which is particularly beneficial in noisy or heterogeneous datasets.

A complete learning pipeline was developed, including IT2-based fuzzification, data-driven hybrid rule generation, dominance-based rule weighting, and ant colony optimization for compact and high-quality rule subset selection. The dominance framework incorporates both fuzzy dominance, reflecting linguistic relevance, and error-based dominance, capturing numerical predictive performance, thereby ensuring that selected rules are both interpretable and accurate. The inference mechanism employs simplified type reduction and constrained TSK outputs to maintain semantic consistency between antecedents and consequents while ensuring computational efficiency during prediction.

Extensive experimental evaluation on benchmark regression datasets demonstrated that the proposed HIT2-MTSK framework consistently outperforms classical fuzzy modeling approaches and achieves competitive performance relative to non-fuzzy machine learning models. The results confirm that hybridization of Mamdani and TSK paradigms within an IT2 fuzzy framework can effectively mitigate the traditional trade-off between explainability and accuracy. Overall, this chapter establishes HIT2-MTSK as a robust and interpretable regression model and provides a foundation for subsequent chapters that extend the framework toward uncertainty quantification, adaptive learning, and real-world deployment scenarios.

# Chapter 6

# Conclusion

## 6.1 Research Summary

This thesis has explored advancements in fuzzy logic systems, focusing on enhancing explainability while maintaining or improving predictive accuracy in time-series classification and regression tasks. The research has also explored semantic consistency in fuzzy regression systems. The research was motivated by the persistent trade-off between interpretability and performance in fuzzy systems, particularly in domains requiring transparent decision-making.

The introductory chapter established the research context, highlighting the limitations of traditional fuzzy approaches and the need for innovations in temporal modelling, semantic consistency, and hybrid inference mechanisms. The background chapter contextualised existing theoretical frameworks and methodological approaches, identifying gaps in knowledge and application in our areas of interest.

In Chapter 3, the Optimized Time Series Interval Type-2 Fuzzy System (OTS-IVFS) was presented to address challenges in explainable time-series classification. By introducing automated time period discovery and linguistic temporal features, the approach demonstrated competitive accuracy on benchmark datasets while providing interpretable rules aligned with human temporal reasoning. Experimental results showed substantial improvements in domains like seismic prediction, where the method outperformed baselines by up to 16.67% in accuracy.

The logical gaps chapter formalised inconsistencies in fuzzy regression systems where predictions violate known monotonic relationships. A gap detection algorithm and targeted rule insertion mechanism were proposed, successfully restoring semantic consistency in real-world financial datasets with minimal impact on predictive performance. The repair process preserved model transparency, ensuring that augmented rule bases remained interpretable.

Finally, the hybrid regression chapter presented a novel Hybrid Interval Type-2 Mamdani-TSK (HIT2-MTSK) framework that integrates linguistic interpretability with numerical precision through dual-component rules and constrained polynomials. Using ant colony optimisation and dual weights, the system achieved superior results on KEEL benchmarks, outperforming pure fuzzy methods in four out of six datasets while maintaining full explainability.

Collectively, these chapters advance fuzzy logic towards more trustworthy systems, demonstrating that enhanced explainability need not compromise accuracy in predictive modelling.

## 6.2 Research Contributions

This thesis makes several novel contributions to fuzzy logic and explainable AI, spanning methodological innovations, algorithmic developments, and empirical insights. The contributions are detailed below, organised by research area for clarity:

- **Time-Series Classification:**
  - Developed the Optimized Time Series Interval Type-2 Fuzzy System (OTS-IVFS), introducing correlation-based time step selection with polynomial fitting for automated identification of predictive temporal patterns.
  - Proposed optimised fuzzy set generation over time periods, enabling linguistic features (e.g., “recent” or “long-term”) that enhance model interpretability while providing noise-robust representations.
  - Demonstrated empirical superiority over traditional fuzzy time-series methods, with accuracy improvements up to 16.67% on benchmark datasets, whilst maintaining semantic clarity in temporal rules.
- **Logical Gaps in Fuzzy Regression:**

  - Formalised the concept of semantic inconsistencies in fuzzy rule bases, particularly violations of monotonic relationships in regression tasks.
  - Introduced a detection framework for systematic identification of logical gaps in rule base coverage.
  - Proposed a repair algorithm using quality metrics for targeted rule insertion, showing that small interventions restore consistency with minimal impact on predictive accuracy and rule base size.

- **Hybrid Regression Modelling:**
  - Created a novel dual-component rule structure combining Mamdani linguistic consequents with constrained TSK polynomials for balanced interpretability and precision.
  - Developed a dual-weight mechanism (fuzzy support-confidence and error-based) integrated with ant colony optimisation for enhanced rule selection and inference.
  - Achieved RMSE improvements of 0.4% to 19% over fuzzy baselines on KEEL datasets, whilst preserving full explainability and outperforming non-fuzzy methods in selected cases.

These contributions collectively extend the theoretical foundations of fuzzy systems, providing reusable methodologies that enhance explainability across classification and regression tasks. The work also contributes empirical evidence from diverse datasets, highlighting the practical value of interval type-2 fuzzy sets and evolutionary techniques in real-world applications.

## 6.3 Potential Applications

The methodologies developed in this thesis have broad applicability across domains requiring explainable predictive modelling.

In financial sectors, OTS-IVFS could enhance liquidity prediction systems where linguistic temporal rules (e.g., "recent market volatility", "medium-term price trends") provide risk managers with interpretable insights whilst maintaining predictive accuracy. The HIT2-MTSK framework proves particularly valuable for credit scoring and risk assessment, combining precise predictions with auditable rules to meet regulatory requirements. The logical gap repair algorithm ensures monotonic relationships in pricing models, critical for maintaining fairness and preventing discriminatory lending practices.

Industrial control systems represent another key area, where the hybrid model's precision supports process optimisation whilst explainable rules enable operator oversight. Manufacturing plants could employ these methods for predictive maintenance, with temporal features capturing patterns like "recent vibration increase" or "long-term degradation trends". The semantic consistency enforcement prevents illogical control decisions that could compromise safety or efficiency.

Environmental monitoring and climate modelling benefit from the interpretable temporal patterns identified by OTS-IVFS. Seismic monitoring, as demonstrated in evaluations, could employ the framework for earthquake prediction with transparent temporal features. Weather forecasting systems could leverage linguistic rules to communicate predictions more effectively to the public, whilst maintaining the accuracy needed for critical decisions.

In smart city applications, the methodologies enable transparent decision-making for traffic management, energy distribution, and resource allocation. The time-series approach could identify patterns in urban mobility, whilst the hybrid regression framework optimises resource utilisation with explainable policies. Transportation networks could benefit from interpretable models that justify routing decisions and predict congestion patterns using human-understandable temporal concepts.

Theoretically, these approaches contribute to explainable AI frameworks, potentially integrating with broader systems for causal reasoning or uncertainty quantification. In regulated industries like insurance or autonomous vehicles, the emphasis on semantic consistency and interpretability aligns with emerging standards for trustworthy AI.

Overall, the research enables deployment in high-stakes environments where transparency is paramount, potentially reducing adoption barriers for fuzzy systems in practical settings.

## 6.4 Future Work

Several promising research directions emerge from this work, spanning methodological, theoretical, and applied aspects of fuzzy systems for explainable predictive modelling. These directions aim to further enhance adaptability, scalability, robustness, and real-world applicability of the proposed approaches while preserving semantic interpretability.

### 6.4.1 Adaptive and Online Learning for Time-Series Models

While the current time-series framework focuses on offline learning with static datasets, many real-world environments exhibit non-stationary behaviour and concept drift. Future work could investigate online and incremental learning mechanisms for fuzzy time-series classifiers, allowing rule bases and membership functions to evolve as new data arrive. Adaptive windowing strategies and dynamic rule pruning could help maintain model compactness while tracking changing temporal patterns. Integrating drift detection techniques could further improve robustness in long-term deployments, particularly in domains such as finance, healthcare monitoring, and industrial diagnostics.

### 6.4.2 Automated and Multivariate Temporal Pattern Discovery

The present time-series methodology primarily addresses univariate temporal dependencies. Extending the framework to multivariate time-series would allow modelling of interactions among multiple correlated signals. Automated discovery of cross-variable temporal relationships and lag structures could be incorporated into the linguistic rule generation process. This would support more expressive temporal rules while retaining interpretability, enabling applications in sensor fusion, environmental monitoring, and physiological signal analysis.

### 6.4.3 Enhanced Semantic Consistency and Rule Base Integrity

Although semantic consistency was addressed in fuzzy regression, future research could explore stronger logical guarantees during rule base construction. Preventive constraint mechanisms could be embedded directly into optimisation algorithms to avoid generating contradictory or redundant rules. Extensions to handle non-monotonic relationships, ordinal outputs, and classification tasks would broaden applicability beyond regression settings. Moreover, formal verification techniques from rule-based systems could be adapted to certify consistency, completeness, and coverage of fuzzy rule bases.

### 6.4.4 Advanced Optimisation and Learning Strategies

The thesis primarily employs metaheuristic optimisation methods such as Ant Colony Optimisation for rule selection. Future studies could explore hybrid optimisation schemes combining global metaheuristics with local gradient-based fine-tuning of membership function parameters and TSK consequent coefficients. Multi-objective optimisation frameworks could explicitly trade off prediction accuracy, interpretability, rule base size, and computational complexity. This would provide principled control over explainability–performance trade-offs tailored to application-specific requirements.

### 6.4.5 Higher-Order and Deep Neuro-Fuzzy Hybrid Architectures

The HIT2-MTSK framework demonstrates that combining Mamdani linguistic components with TSK numerical precision yields strong performance with interpretability. Future work could investigate hierarchical or multi-layer hybrid fuzzy architectures, enabling representation of more complex nonlinear mappings while preserving local interpretability. Integration with neural representation learning could support automatic feature extraction from high-dimensional inputs, followed by fuzzy inference layers responsible for interpretable decision-making, forming explainable neuro-fuzzy pipelines.

### 6.4.6 Uncertainty Quantification and Decision Support Integration

Interval Type-2 fuzzy sets already provide a mechanism for modelling uncertainty, yet further work is needed to propagate and quantify predictive uncertainty throughout hybrid inference pipelines. Developing calibrated confidence measures alongside predictions could enhance decision support capabilities in safety-critical systems. Coupling fuzzy outputs with risk-aware decision-making frameworks would improve usability in domains such as medical diagnosis, energy systems, and autonomous control.

### 6.4.7 Scalability, Efficiency, and Hardware Acceleration

Although the proposed models perform well on benchmark datasets, future research should evaluate scalability on large-scale and high-frequency data streams. Parallel implementations and GPU acceleration for fuzzification, inference, and optimisation stages could significantly reduce training and inference times. Additionally, investigating lightweight implementations for edge devices would enable deployment in embedded and Internet-of-Things environments where explainability and low-power operation are both critical.

### 6.4.8 Human-in-the-Loop and Interactive Explainability

Interpretability is most valuable when aligned with human understanding and domain expertise. Future work could incorporate human-in-the-loop mechanisms allowing experts to inspect, modify, and validate linguistic rules during training. Interactive visual analytics tools could support exploration of fuzzy partitions, temporal patterns, and rule activations, improving trust, debugging, and collaborative model development in applied settings.

### 6.4.9 Fairness, Ethics, and Regulatory Compliance

As explainable AI becomes increasingly regulated, future extensions should assess fairness and bias within fuzzy inference systems. Investigating how linguistic rules and membership functions may encode or mitigate bias is essential for deployment in sensitive domains. Formal links between fuzzy explanations and regulatory requirements such as transparency, contestability, and auditability could further strengthen adoption in real-world governance frameworks.

### 6.4.10 Benchmarking and Standardised Evaluation Protocols

Finally, broader benchmarking across diverse datasets and problem domains would facilitate more rigorous evaluation of explainable fuzzy systems. Developing standardised metrics that jointly assess accuracy, interpretability, robustness, and uncertainty representation would allow fairer comparison with both black-box machine learning models and alternative interpretable approaches.

Collectively, these future research directions aim to advance fuzzy systems toward more adaptive, scalable, and trustworthy intelligent systems. By strengthening both predictive capability and human interpretability, future developments can further bridge the gap between data-driven learning and transparent decision-making in complex real-world environments.

# Bibliography


[1] L. A. Zadeh, "Fuzzy sets," *Information and control*, vol. 8, no. 3, pp. 338–353, 1965, publisher: Elsevier. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S001999586590241X

[2] O. Cordón, "A historical review of Mamdani-type genetic fuzzy systems," in *Combining Experimentation and Theory – A Hommage to Abe Mamdani*, ser. Studies in Fuzziness and Soft Computing, E. Trillas, P. P. Bonissone, L. Magdalena, and J. Kacprzyk, Eds. Berlin, Heidelberg: Springer, 2012, vol. 271, pp. 73–90.

[3] J. G. De Gooijer and R. J. Hyndman, "25 Years of Time Series Forecasting," *International Journal of Forecasting*, vol. 22, no. 3, pp. 443–473, 2006.

[4] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "The M4 Competition: 100,000 time series and 61 forecasting methods," *International Journal of Forecasting*, vol. 36, no. 1, pp. 54–74, 2020.

[5] O. B. Sezer, M. U. Gudelek, and A. M. Ozbayoglu, "Financial time series forecasting with deep learning: A systematic literature review: 2005–2019," *Applied Soft Computing Journal*, vol. 90, pp. 2005–2019, 2020.

[6] J. Chen, L. Song, M. J. Wainwright, and M. I. Jordan, "Learning to Explain: An Information-Theoretic Perspective on Model Interpretation," *Proceedings of Machine Learning Research*, vol. 80, pp. 883–892, 2018, publisher: PMLR.

[7] B. Lim, S. O. Arık, N. Loeff, and T. Pfister, "Temporal Fusion Transformers for interpretable multi-horizon time series forecasting," *International Journal of Forecasting*, vol. 37, no. 4, pp. 1748–1764, 2021.

[8] H. Zhou, S. Zhang, J. Peng, S. Zhang, J. Li, H. Xiong, and W. Zhang, "Informer: Beyond efficient transformer for long sequence time-series forecasting," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, pp. 11 106–11 115, 2021.

[9] M. Middlehurst, P. Schäfer, and A. Bagnall, "Bake off redux: a review and experimental evaluation of recent time series classification algorithms," *Data Mining and Knowledge Discovery*, vol. 38, no. 4, pp. 1959–2028, 2024.

[10] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 1798–1828, 2013.

[11] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in Neural Information Processing Systems*, pp. 5998–6008, 2017.

[12] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, publisher: Nature Publishing Group.

[13] E. Strubell, A. Ganesh, and A. McCallum, "Energy and policy considerations for deep learning in NLP," *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pp. 3645–3650, 2019.

[14] R. Schwartz, J. Dodge, N. A. Smith, and O. Etzioni, "Green AI," *Communications of the ACM*, vol. 63, no. 12, pp. 54–63, 2020.

[15] A. Rajkomar, E. Oren, K. Chen, A. M. Dai, N. Hajaj, M. Hardt, P. J. Liu, X. Liu, J. Marcus, M. Sun, P. Sundberg, H. Yee, K. Zhang, Y. Zhang, G. Flores, G. E. Duggan, J. Irvine, Q. Le, K. Litsch, A. Mossin, J. Tansuwan, D. Wang, J. Wexler, J. Wilson, D. Ludwig, S. L. Volchenboum, K. Chou, M. Pearson, S. Madabushi, N. H. Shah, A. J. Butte, M. D. Howell, C. Cui, G. S. Corrado, and J. Dean, "Scalable and accurate deep learning with electronic health records," *npj Digital Medicine*, vol. 1, no. 1, p. 18, 2018.

[16] E. J. Topol, "High-performance medicine: the convergence of human and artificial intelligence," *Nature Medicine*, vol. 25, no. 1, pp. 44–56, 2019.

[17] C. Rudin, "Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead," *Nature Machine Intelligence*, vol. 1, no. 5, pp. 206–215, 2019, publisher: Nature Publishing Group.

[18] C. Molnar, *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*, 2nd ed. Lulu.com, 2020.

[19] R. Caruana, Y. Lou, J. Gehrke, P. Koch, M. Sturm, and N. Elhadad, "Intelligible models for healthcare: Predicting pneumonia risk and hospital 30-day readmission," *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1721–1730, 2015.

[20] J. Angwin, J. Larson, S. Mattu, and L. Kirchner, "Machine bias: There's software used across the country to predict future criminals. And it's biased against blacks," *ProPublica*, 2016.

[21] A. Karpatne, G. Atluri, J. H. Faghmous, M. Steinbach, A. Banerjee, A. Ganguly, S. Shekhar, N. Samatova, and V. Kumar, "Theory-guided data science: A new paradigm for scientific discovery from data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, no. 10, pp. 2318–2331, 2017.

[22] J. Willard, X. Jia, S. Xu, M. Steinbach, and V. Kumar, "Integrating physics-based modeling with machine learning: A survey," *arXiv preprint arXiv:2003.04919*, 2021.

[23] G. Marcus and E. Davis, *Rebooting AI: Building artificial intelligence we can trust*. Pantheon Books, 2020.

[24] B. M. Lake, T. D. Ullman, J. B. Tenenbaum, and S. J. Gershman, "Building machines that learn and think like people," *Behavioral and Brain Sciences*, vol. 40, no. 2, p. e253, 2017.

[25] A. Bibal, M. Lognoul, A. de Streel, and B. Frénay, "Legal requirements on explainability in machine learning," *Artificial Intelligence and Law*, vol. 29, no. 2, pp. 149–169, 2021.

[26] European Union, "General Data Protection Regulation," 2016.

[27] E. Tjoa and C. Guan, "A survey on explainable artificial intelligence (XAI): Toward medical XAI," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 11, pp. 4793–4813, 2021.

[28] A. Adadi and M. Berrada, "Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)," *IEEE Access*, vol. 6, pp. 52 138–52 160, 2018, publisher: IEEE.

[29] E. H. Mamdani and S. Assilian, "An experiment in linguistic synthesis with a fuzzy logic controller," *International journal of man-machine studies*, vol. 7, no. 1, pp. 1–13, 1975, publisher: Elsevier. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0020737375800022

[30] M.-Y. Chen and B.-T. Chen, "A hybrid fuzzy time series model based on granular computing for stock price forecasting," *Information Sciences*, vol. 294, pp. 227–241, 2015, publisher: Elsevier. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0020025514009505

[31] H. J. Sadaei, P. C. de Lima e Silva, F. G. Guimarães, and M. H. Lee, "Short-term load forecasting by using a combined method of convolutional neural networks and fuzzy time series," *Energy*, vol. 175, pp. 365–377, 2019.

[32] Q. Song and B. S. Chissom, "Fuzzy time series and its models," *Fuzzy sets and systems*, vol. 54, no. 3, pp. 269–277, 1993, publisher: Elsevier. [Online]. Available: https://www.sciencedirect.com/science/article/pii/016501149390372O

[33] F. Gaxiola, P. Melin, F. Valdez, and O. Castillo, "Interval type-2 fuzzy weight adjustment for backpropagation neural networks with application in time series prediction," *Information Sciences*, vol. 260, pp. 1–14, Mar. 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0020025513007937

[34] A. Jastrzebska, G. Nápoles, W. Homenda, and K. Vanhoof, "Fuzzy Cognitive Map-Driven Comprehensive Time-Series Classification," *IEEE Transactions on Cybernetics*, vol. 53, no. 2, pp. 1348–1359, Feb. 2023. [Online]. Available: https://ieeexplore.ieee.org/document/9660779

[35] W.-H. Au and K. Chan, "Mining fuzzy rules for time series classification," in *2004 IEEE International Conference on Fuzzy Systems*, vol. 1, Jul. 2004, pp. 239–244 vol.1, iSSN: 1098-7584. [Online]. Available: https://ieeexplore.ieee.org/document/1375726

[36] V. Novák, I. Perfilieva, M. Holčapek, and V. Kreinovich, "Filtering out high frequencies in time series using F-transform," *Information Sciences*, vol. 274, pp. 192–209, 2014, publisher: Elsevier. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0020025514002394

[37] G. Nápoles, Y. Salgueiro, I. Grau, and M. L. Espinosa, "Recurrence-aware long-term cognitive network for explainable pattern classification," *IEEE transactions on cybernetics*, vol. 53, no. 10, pp. 6083–6094, 2022, publisher: IEEE. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9764661/

[38] M. Latinovic, I. Dragovic, V. Arsić, and B. Petrovic, "A Fuzzy Inference System for Credit Scoring Using Boolean Consistent Fuzzy Logic," *International Journal of Computational Intelligence Systems*, vol. 11, no. 1, pp. 414–427, 2018.

[39] G. Bennouna and M. Tkiouat, "Fuzzy Logic Approach Applied to Credit Scoring for Microfinance in Morocco," *Procedia Computer Science*, vol. 127, pp. 274–283, 2018.

[40] M. Sanchez-Roger, M. D. Oliver-Alfonso, and C. Sanchís-Pedregosa, "Fuzzy Logic and Its Uses in Finance: A Systematic Review Exploring Its Potential to Deal with Banking Crises," *Mathematics*, vol. 7, no. 11, p. 1091, 2019.

[41] P. Shinde and S. Shah, "A review of machine learning and deep learning applications," in *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*. Institute of Electrical and Electronics Engineers Inc, 2018, pp. 1–6.

[42] B. Goodman and S. Flaxman, "European Union Regulations on Algorithmic Decision-Making and a "Right to Explanation"," *AI Magazine*, vol. 38, no. 3, pp. 50–57, 2017, publisher: AAAI.

[43] S. Wachter, B. Mittelstadt, and L. Floridi, "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation," *Interna-*

*tional Data Privacy Law*, vol. 7, no. 2, pp. 76–99, 2017, publisher: Oxford University Press.

[44] T. Hastie, R. Tibshirani, and J. H. Friedman, *The elements of statistical learning: data mining, inference, and prediction*, 2nd ed., ser. Springer Series in Statistics. New York, NY: Springer, 2017.

[45] D. C. Montgomery, E. A. Peck, and G. G. Vining, *Introduction to Linear Regression Analysis*, 6th ed. Wiley, 2021.

[46] F. Maturo and S. Hoskova-Mayerova, "Fuzzy Regression Models and Alternative Operations for Economic and Social Sciences," in *Recent Trends in Social Systems: Quantitative Theories and Quantitative Models*, ser. Studies in Systems, Decision and Control. Springer, 2017, vol. 66, pp. 235–247.

[47] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An introduction to statistical learning: with applications in R*, 2nd ed., ser. Springer texts in statistics. New York, NY: Springer, 2021.

[48] R. F. Engle, "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, vol. 50, no. 4, pp. 987–1007, 1982, publisher: The Econometric Society.

[49] H. Motulsky and A. Christopoulos, *Fitting models to biological data using linear and nonlinear regression: a practical guide to curve fitting*. Oxford University Press, 2004. [Online]. Available: https://books.google.com/books?hl=en&lr=&id=g1FO9pquF3kC&oi=fnd&pg=PA4&dq=motulsky+fitting&ots=m3S6egcF5p&sig=DKSR7CTuGbx1DGweynUtYS853FI

[50] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-Based Learning Applied to Document Recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998, publisher: IEEE.

[51] Federal Deposit Insurance Corporation, "Credit Risk Grading Systems: Observations from a Horizontal Assessment," FDIC, Tech.

Rep., 2018. [Online]. Available: https://www.fdic.gov/bank-examinations/credit-risk-grading-systems-observations-horizontal-assessment

[52] G. C. Weston, M. J. Rothe, B. D. Kels, and J. M. Grant-Kels, "The good, the bad, and the ugly of medication coverage: Is altering a diagnosis to ensure medication coverage ethical?" *International Journal of Women's Dermatology*, vol. 2, no. 2, pp. 67–68, 2016, publisher: Elsevier.

[53] W. J. von Eschenbach, "Transparency and the Black Box Problem: Why We Do Not Trust AI," *Philosophy & Technology*, vol. 34, no. 4, pp. 1607–1622, Dec. 2021. [Online]. Available: https://doi.org/10.1007/s13347-021-00477-0

[54] I. Rather, S. Kumar, and A. Gandomi, "Breaking the data barrier: a review of deep learning techniques for democratizing AI with small datasets," *Artificial Intelligence Review*, vol. 57, 2024.

[55] W. Saeed and C. Omlin, "Explainable AI (XAI): A systematic meta-survey of current challenges and future opportunities," *Knowledge-based systems*, vol. 263, p. 110273, 2023, publisher: Elsevier. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0950705123000230

[56] A. B. Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, S. Garcia, S. Gil-Lopez, D. Molina, R. Benjamins, R. Chatila, and F. Herrera, "Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI," *Information Fusion*, vol. 58, pp. 82–115, 2020, publisher: Elsevier.

[57] L. A. Zadeh, "The birth and evolution of fuzzy logic," *International Journal of General System*, vol. 17, no. 2-3, pp. 95–105, 1990.

[58] S. Guillaume and L. Magdalena, "Expert Guided Integration of Induced Knowledge into a Fuzzy Knowledge Base," *Soft Computing*, vol. 10, no. 8, pp. 773–784, 2006, publisher: Springer.

[59] L. Zadeh, "Is there a need for fuzzy logic?" *Information Sciences*, vol. 178, pp. 2751–2779, 2008.

[60] J. M. Alonso, C. Castiello, and C. Mencar, “Interpretability of Fuzzy Systems: Current Research Trends and Prospects,” *IEEE Computational Intelligence Magazine*, vol. 16, no. 1, pp. 14–28, 2021, publisher: IEEE.

[61] T. Takagi and M. Sugeno, “Fuzzy identification of systems and its applications to modeling and control,” *IEEE Transactions on Systems, Man and Cybernetics*, vol. SMC-15, pp. 116–132, 1985.

[62] O. Cordón, F. Herrera, and P. Villar, “Generating the knowledge base of a fuzzy rule-based system by the genetic learning of the data base,” *IEEE Transactions on fuzzy systems*, vol. 9, no. 4, pp. 667–674, 2001, publisher: IEEE. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/940977/

[63] M. J. Gacto, R. Alcalá, and F. Herrera, “Interpretability of Linguistic Fuzzy Rule-Based Systems: An Overview of Interpretability Measures,” *Information Sciences*, vol. 181, no. 20, pp. 4340–4360, 2011.

[64] P. P. Angelov, D. P. Filev, and N. Kasabov, “Evolving Intelligent Systems: Methodology and Applications,” *IEEE Transactions on Fuzzy Systems*, vol. 19, no. 2, pp. 201–211, 2011, publisher: IEEE.

[65] D. Wu and W. W. Tan, “On the Functional Equivalence of TSK Fuzzy Systems to Neural Networks, Mixture of Experts, CART, and Stacking Ensemble Regression,” *IEEE Transactions on Fuzzy Systems*, vol. 28, no. 10, pp. 2570–2580, 2019, publisher: IEEE.

[66] J. Yen, L. Wang, and C. Gillespie, “Improving the interpretability of tsk fuzzy models by combining global learning and local learning,” *IEEE Transactions on Fuzzy Systems*, vol. 6, pp. 530–537, 1998.

[67] A. Holzinger, G. Langs, H. Denk, K. Zatloukal, and H. Müller, “Causability and Explainability of Artificial Intelligence in Medicine,” *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 9, no. 4, p. e1312, 2019, publisher: Wiley.

[68] P. Bracke, A. Datta, C. Jung, and S. Sen, “Machine Learning Explainability in Finance: An Application to Default Risk Analysis,” Bank of England, Staff Working Paper 816, 2019.

[69] A. K. Varshney and V. Torra, “Literature Review of the Recent Trends and Applications in Various Fuzzy Rule-Based Systems,” *International Journal of Fuzzy Systems*, vol. 25, no. 6, pp. 2163–2186, Sep. 2023. [Online]. Available: https://link.springer.com/10.1007/s40815-023-01534-w

[70] E. Mamdani, “Application of fuzzy algorithms for control of simple dynamic plant,” *Proceedings of the Institution of Electrical Engineers*, vol. 121, no. 12, p. 1585, 1974. [Online]. Available: https://digital-library.theiet.org/content/journals/10.1049/piee.1974.0328

[71] L. A. Zadeh, “The concept of a linguistic variable and its application to approximate reasoning-III,” *Information sciences*, vol. 9, no. 1, pp. 43–80, 1975, publisher: Elsevier. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0020025575900171

[72] J. Mendel, *Uncertain Rule-Based Fuzzy Systems: Introduction and New Directions*, 2nd ed. Springer, 2017.

[73] H. Hagras, “Type-2 FLCs: A new generation of fuzzy controllers,” *IEEE Computational Intelligence Magazine*, vol. 2, no. 1, pp. 30–43, 2007.

[74] N. Karnik, J. Mendel, and Q. Liang, “Type-2 fuzzy logic systems,” *IEEE Transactions on Fuzzy Systems*, vol. 7, pp. 643–658, 1999.

[75] D. Wu and J. M. Mendel, “Enhanced karnik–mendel algorithms,” *IEEE transactions on fuzzy systems*, vol. 17, no. 4, pp. 923–934, 2008, publisher: IEEE. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/4505357/

[76] M. Nie and W. W. Tan, “Towards an efficient type-reduction method for interval type-2 fuzzy logic systems,” in *2008 IEEE International Conference on Fuzzy Systems (IEEE World Congress on Computational Intelligence)*. IEEE, 2008, pp. 1425–1432. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/4630559/

[77] H. Wu and J. M. Mendel, "Uncertainty bounds and their use in the design of interval type-2 fuzzy logic systems," *IEEE Transactions on fuzzy systems*, vol. 10, no. 5, pp. 622–639, 2002, publisher: IEEE. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/1038818/

[78] H. Ying, "Sufficient conditions on uniform approximation of multivariate functions by general Takagi-Sugeno fuzzy systems with linear rule consequent," *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 28, no. 4, pp. 515–520, 1998, publisher: IEEE. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/686713/

[79] Q. Liang and J. Mendel, "Introduction to type-2 tsk fuzzy logic systems," in *1999 IEEE International Conference on Fuzzy Systems*, vol. 3, 1999, pp. 1534–1539.

[80] L. Pickering, K. Cohen, and B. De Baets, "A Narrative Review on the Interpretability of Fuzzy Rule-Based Models from a Modern Interpretable Machine Learning Perspective," *International Journal of Fuzzy Systems*, Mar. 2025. [Online]. Available: https://doi.org/10.1007/s40815-025-02022-z

[81] T. Bäck and H.-P. Schwefel, "An overview of evolutionary algorithms for parameter optimization," *Evolutionary computation*, vol. 1, no. 1, pp. 1–23, 1993, publisher: mit Press. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/6791438/

[82] J. Koza, "Genetic programming as a means for programming computers by natural selection," *Statistics and Computing*, vol. 4, no. 2, Jun. 1994. [Online]. Available: http://link.springer.com/10.1007/BF00175355

[83] P. R. Thrift, "Fuzzy Logic Synthesis with Genetic Algorithms." in *ICGA*, 1991, pp. 509–513. [Online]. Available: https://www.academia.edu/download/36706667/1991-Thrift-ICGA91.pdf

[84] H. Ishibuchi, K. Nozaki, N. Yamamoto, and H. Tanaka, "Selecting fuzzy if-then rules for classification problems using genetic algorithms," *IEEE Transactions on fuzzy systems*, vol. 3, no. 3, pp. 260–270, 1995, publisher: IEEE. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/413232/

[85] F. Santoso, M. A. Garratt, and S. G. Anavatti, “T2-ETS-IE: A type-2 evolutionary Takagi–Sugeno fuzzy inference system with the information entropy-based pruning technique,” *IEEE transactions on fuzzy systems*, vol. 28, no. 10, pp. 2665–2672, 2019, publisher: IEEE. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8848600/

[86] W.-C. Su, C.-F. Juang, and C.-M. Hsu, “Multiobjective evolutionary interpretable type-2 fuzzy systems with structure and parameter learning for hexapod robot control,” *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 52, no. 5, pp. 3066–3078, 2021, publisher: IEEE. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9380732/

[87] P. Melin, I. Miramontes, and G. Prado-Arechiga, “A hybrid model based on modular neural networks and fuzzy systems for classification of blood pressure and hypertension risk diagnosis,” *Expert Systems with Applications*, vol. 107, pp. 146–164, 2018, publisher: Elsevier. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S095741741830263X

[88] J. T. Bishop, M. Gallagher, and W. N. Browne, “A genetic fuzzy system for interpretable and parsimonious reinforcement learning policies,” in *Proceedings of the Genetic and Evolutionary Computation Conference Companion*. Lille France: ACM, Jul. 2021, pp. 1630–1638. [Online]. Available: https://dl.acm.org/doi/10.1145/3449726.3463198

[89] X. Gu, J. Han, Q. Shen, and P. P. Angelov, “Autonomous learning for fuzzy systems: a review,” *Artificial Intelligence Review*, vol. 56, no. 8, pp. 7549–7595, Aug. 2023. [Online]. Available: https://doi.org/10.1007/s10462-022-10355-6

[90] X. Gu, Q. Shen, and P. Angelov, “Particle swarm optimized autonomous learning fuzzy system,” *IEEE Transactions on Cybernetics*, vol. 51, no. 11, pp. 5352–5363, 2021.

[91] X. Gu and Q. Shen, “A self-adaptive fuzzy learning system for streaming data prediction,” *Information Sciences*, vol. 579, pp. 623–647, 2021.

[92] J. Mendel, H. Hagras, W.-W. Tan, W. W. Melek, and H. Ying, *Introduction to type-2 fuzzy logic control: theory and applications*. John Wiley & Sons, 2014. [Online]. Available: https://books.google.com/books?hl=en&lr=&id=C1HcAwAAQBAJ&oi=fnd&pg=PT10&dq=mendel+2014+type-2&ots=uUUuLox7Ne&sig=wAAHYVRTGOldmW720snSRZyRFxA

[93] X. Gu, P. Angelov, and Q. Shen, "Self-organising fuzzy belief inference system for classification," *IEEE Transactions on Fuzzy Systems*, vol. 30, no. 12, pp. 5473–5483, 2022.

[94] H. Ishibuchi, "Genetic fuzzy systems: evolutionary tuning and learning of fuzzy knowledge bases," *Fuzzy Sets and Systems*, vol. 141, pp. 1–3, 2004.

[95] X. Gu, P. Angelov, J. Han, and Q. Shen, "Multilayer evolving fuzzy neural networks," *IEEE Transactions on Fuzzy Systems*, vol. 31, no. 12, pp. 4158–4169, 2023.

[96] X. Gu, Q. Ni, and Q. Shen, "Multilayer evolving fuzzy neural networks with self-adaptive dimensionality compression for high-dimensional data classification," *IEEE Transactions on Fuzzy Systems*, vol. 32, no. 11, pp. 6314–6328, 2024.

[97] O. Cordón, F. Gomide, F. Herrera, F. Hoffmann, and L. Magdalena, "Ten years of genetic fuzzy systems: current framework and new trends," *Fuzzy Sets and Systems*, vol. 141, no. 1, pp. 5–31, Jan. 2004, publisher: Elsevier BV. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0165011403001118

[98] S. Guillaume, "Designing fuzzy inference systems from data: An interpretability-oriented review," *IEEE Transactions on fuzzy systems*, vol. 9, no. 3, pp. 426–443, 2002, publisher: IEEE. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/928739/

[99] S. Izquierdo and L. R. Izquierdo, "Mamdani fuzzy systems for modelling and simulation: A critical assessment," *Available at SSRN 2900827*, 2017. [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2900827

[100] Q. Song and B. S. Chissom, "Forecasting enrollments with fuzzy time series—Part I," *Fuzzy sets and systems*, vol. 54, no. 1, pp. 1–9, 1993, publisher: Elsevier. [Online]. Available: https://www.sciencedirect.com/science/article/pii/016501149390355L

[101] ——, "Forecasting enrollments with fuzzy time series—Part II," *Fuzzy sets and systems*, vol. 62, no. 1, pp. 1–8, 1994, publisher: Elsevier. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0165011494900671

[102] S.-M. Chen, "Forecasting enrollments based on fuzzy time series," *Fuzzy sets and systems*, vol. 81, no. 3, pp. 311–319, 1996, publisher: Elsevier. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0165011495002200

[103] K. Huarng, "Effective lengths of intervals to improve forecasting in fuzzy time series," *Fuzzy sets and systems*, vol. 123, no. 3, pp. 387–394, 2001, publisher: Elsevier. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0165011400000579

[104] R.-C. Tsaur, J.-C. O. Yang, and H.-F. Wang, "Fuzzy relation analysis in fuzzy time series model," *Computers & Mathematics with Applications*, vol. 49, no. 4, pp. 539–548, 2005, publisher: Elsevier. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0898122105000489

[105] H.-K. Yu, "Weighted fuzzy time series models for TAIEX forecasting," *Physica A: Statistical Mechanics and its Applications*, vol. 349, no. 3-4, pp. 609–624, 2005, publisher: Elsevier. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378437104014128

[106] K. Huarng, "Heuristic models of fuzzy time series for forecasting," *Fuzzy sets and systems*, vol. 123, no. 3, pp. 369–386, 2001, publisher: Elsevier. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0165011400000932

[107] E. Egrioglu, C. H. Aladag, and U. Yolcu, "Fuzzy time series forecasting with a novel hybrid approach combining fuzzy c-means and neural networks," *Expert Systems with Applications*, vol. 40, no. 3, pp. 854–857, 2013, publisher: Elsevier. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417412007567

[108] S. Panigrahi and H. S. Behera, “A study on leading machine learning techniques for high order fuzzy time series forecasting,” *Engineering applications of artificial intelligence*, vol. 87, p. 103245, 2020, publisher: Elsevier. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S095219761930226X

[109] I.-H. Kuo, S.-J. Horng, T.-W. Kao, T.-L. Lin, C.-L. Lee, and Y. Pan, “An improved method for forecasting enrollments based on fuzzy time series and particle swarm optimization,” *Expert systems with applications*, vol. 36, no. 3, pp. 6108–6117, 2009, publisher: Elsevier. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417408004818

[110] N. Kumar and S. Susan, “Non-stationary fuzzy time series modeling and forecasting using deep learning with swarm optimization,” *International Journal of Machine Learning and Cybernetics*, Mar. 2025. [Online]. Available: https://doi.org/10.1007/s13042-025-02585-1

[111] P. C. d. L. e Silva, C. A. S. Junior, M. A. Alves, R. Silva, M. W. Cohen, and F. G. Guimarães, “Forecasting in non-stationary environments with fuzzy time series,” *Applied Soft Computing*, vol. 97, p. 106825, 2020, publisher: Elsevier. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1568494620307638

[112] B. Wang and X. Liu, “Adaptive Non-Stationary Fuzzy Time Series Forecasting with Bayesian Networks,” *Sensors*, vol. 25, no. 5, p. 1628, 2025, publisher: MDPI. [Online]. Available: https://www.mdpi.com/1424-8220/25/5/1628

[113] N. S. Bajestani and A. Zare, “Forecasting TAIEX using improved type 2 fuzzy time series,” *Expert Systems with Applications*, vol. 38, no. 5, pp. 5816–5821, 2011, publisher: Elsevier. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417410011930

[114] L.-W. Lee, L.-H. Wang, S.-M. Chen, and Y.-H. Leu, “Handling forecasting problems based on two-factors high-order fuzzy time series,” *IEEE Transactions on fuzzy Systems*, vol. 14, no. 3, pp. 468–477, 2006, publisher: IEEE. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/1638462/

[115] S.-M. Chen, "Forecasting enrollments based on high-order fuzzy time series," *Cybernetics and Systems*, vol. 33, no. 1, pp. 1–16, Jan. 2002. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/019697202753306479

[116] P. O. Lucas, O. Orang, P. C. Silva, E. M. Mendes, and F. G. Guimaraes, "A tutorial on fuzzy time series forecasting models: Recent advances and challenges," *Learn Nonlinear Models*, vol. 19, pp. 29–50, 2022.

[117] M. Bose and K. Mali, "Designing fuzzy time series forecasting models: A survey," *International Journal of Approximate Reasoning*, vol. 111, pp. 78–99, 2019.

[118] K. Alsalem, "A hybrid time series forecasting approach integrating fuzzy clustering and machine learning for enhanced power consumption prediction," *Scientific Reports*, vol. 15, no. 1, p. 6447, 2025, publisher: Nature Publishing Group UK London. [Online]. Available: https://www.nature.com/articles/s41598-025-91123-8

[119] C. Kocak, E. Egrioglu, and E. Bas, "A new deep intuitionistic fuzzy time series forecasting method based on long short-term memory," *The Journal of Supercomputing*, vol. 77, no. 6, pp. 6178–6196, Jun. 2021, publisher: Springer Science and Business Media LLC. [Online]. Available: https://link.springer.com/10.1007/s11227-020-03503-8

[120] R. Gao and O. Duru, "Parsimonious fuzzy time series modelling," *Expert Systems with Applications*, vol. 156, p. 113447, 2020, publisher: Elsevier. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417420302712

[121] W. J. Friedman, "Time in perspective," *Psychological Bulletin*, vol. 110, no. 4, pp. 652–673, 1990.

[122] M. Middlehurst and A. Bagnall, "The FreshPRINCE: A Simple Transformation Based Pipeline Time Series Classifier," in *Pattern Recognition and Artificial Intelligence*, M. El Yacoubi, E. Granger, P. C. Yuen, U. Pal, and N. Vincent, Eds. Cham: Springer International Publishing, 2022, vol. 13364, pp. 150–161, series Title: Lecture Notes in Computer Science. [Online]. Available: https://link.springer.com/10.1007/978-3-031-09282-4_13

[123] H. Ding, G. Trajcevski, P. Scheuermann, X. Wang, and E. Keogh, "Querying and mining of time series data: experimental comparison of representations and distance measures," *Proceedings of the VLDB Endowment*, vol. 1, no. 2, pp. 1542–1552, 2008.

[124] A. Bagnall and J. Lines, "An experimental evaluation of nearest neighbour time series classification," *arXiv preprint arXiv:1406.4757*, 2014.

[125] H. A. Dau and others, "The UCR time series archive," *IEEE/CAA Journal of Automatica Sinica*, vol. 6, no. 6, pp. 1293–1305, 2019.

[126] A. Bagnall, J. Lines, A. Bostrom, J. Large, and E. Keogh, "The great time series classification bake off: a review and experimental evaluation of recent algorithmic advances," *Data Mining and Knowledge Discovery*, vol. 31, no. 3, pp. 606–660, 2017.

[127] A. P. Ruiz, M. Flynn, J. Large, M. Middlehurst, and A. Bagnall, "The great multivariate time series classification bake off: a review and experimental evaluation of recent algorithmic advances," *Data Mining and Knowledge Discovery*, vol. 35, no. 2, pp. 401–449, 2021.

[128] P. Schäfer and U. Leser, "Multivariate time series classification with WEASEL+MUSE," *arXiv preprint arXiv:1711.11343*, 2017.

[129] M. G. Baydogan and G. Runger, "Time series representation and similarity based on local autopatterns," *Data Mining and Knowledge Discovery*, vol. 30, no. 4, pp. 877–915, 2016.

[130] Y. Wang, R. Emonet, E. Fromont, S. Malinowski, E. Menager, L. Mosser, and R. Tavenard, "Learning interpretable shapelets for time series classification through adversarial regularization," *arXiv preprint arXiv:1906.00917*, 2019.

[131] A. Theissler, F. Spinnato, U. Schlegel, and R. Guidotti, "Explainable AI for time series classification: A review, taxonomy and research directions," *IEEE Access*, vol. 10, pp. 100 700–100 724, 2022.

[132] A. Bonifati, F. Del Buono, F. Guerra, and D. Tiano, "Time2Feat: Learning interpretable representations for multivariate time series clustering," *Proceedings of the VLDB Endowment*, vol. 16, no. 2, pp. 193–201, 2022.

[133] R. Jensen and Q. Shen, "Semantics-preserving dimensionality reduction: Rough and fuzzy-rough approaches," *IEEE Transactions on Knowledge and Data Engineering*, vol. 16, no. 12, pp. 1457–1471, 2004.

[134] L. Zheng, F. Chao, N. Mac Parthain, D. Zhang, and Q. Shen, "Feature grouping and selection: A graph-based approach," *Information Sciences*, vol. 546, pp. 1256–1272, 2021.

[135] W. J. Murdoch, C. Singh, K. Kumbier, R. Abbasi-Asl, and B. Yu, "Definitions, methods, and applications in interpretable machine learning," *Proceedings of the National Academy of Sciences*, vol. 116, no. 44, pp. 22 071–22 080, 2019.

[136] T. Rojat, R. Puget, D. Filliat, J. Del Ser, R. Gelin, and N. Díaz-Rodríguez, "Explainable artificial intelligence (XAI) on timeseries data: A survey," *arXiv preprint arXiv:2104.00950*, 2021.

[137] R. Assaf and A. Schumann, "Explainable deep neural networks for multivariate time series predictions," in *IJCAI International Joint Conference on Artificial Intelligence*, vol. 2019-August, 2019, pp. 6488–6490, issue: 2.

[138] S. A. Siddiqui, D. Mercier, M. Munir, A. Dengel, and S. Ahmed, "TSViz: Demystification of deep learning models for time-series analysis," *IEEE Access*, vol. 7, pp. 67 027–67 040, 2019.

[139] M. Guillemé, V. Masson, L. Rozé, and A. Termier, "Agnostic local explanation for time series classification," in *2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI)*, 2019, pp. 432–439.

[140] H. Meng, C. Wagner, and I. Triguero, "SEGAL time series classification — Stable explanations using a generative model and an adaptive weighting method for LIME," *Neural Networks*, vol. 176, p. 106345, Aug. 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0893608024002697

[141] H. Deng, G. Runger, E. Tuv, and M. Vladimir, "A time series forest for classification and feature extraction," *Information Sciences*, vol. 239, pp. 142–153, 2013.

[142] N. Cabello, E. Naghizade, J. Qi, and L. Kulik, "Fast, accurate and interpretable time series classification through randomization," *arXiv preprint arXiv:2105.14876*, 2021.

[143] A. Dempster, F. Petitjean, and G. I. Webb, "ROCKET: Exceptionally fast and accurate time series classification using random convolutional kernels," *Data Mining and Knowledge Discovery*, vol. 34, no. 5, pp. 1454–1495, 2020.

[144] G. Uribarri, F. Barone, A. Ansuini, and E. Fransén, "Detach-ROCKET: Sequential feature selection for time series classification with random convolutional kernels," *Data Mining and Knowledge Discovery*, 2024.

[145] O. Badi, M. Devanne, A. Ismail-Fawaz, J. Abdullayev, V. Lemaire, S. Berretti, J. Weber, and G. Forestier, "COCALITE: A Hybrid Model COmbining CAtch22 and LITE for Time Series Classification," in *2024 IEEE International Conference on Big Data (BigData)*, Dec. 2024, pp. 1229–1236, iSSN: 2573-2978. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10825872

[146] C. H. Lubba, S. S. Sethi, P. Knaute, S. R. Schultz, B. D. Fulcher, and N. S. Jones, "catch22: CAnonical Time-series CHaracteristics: Selected through highly comparative time-series analysis," *Data Mining and Knowledge Discovery*, vol. 33, no. 6, pp. 1821–1852, Nov. 2019. [Online]. Available: http://link.springer.com/10.1007/s10618-019-00647-x

[147] J. Hills, J. Lines, E. Baranauskas, J. Mapp, and A. Bagnall, "Classification of time series by shapelet transformation," *Data Mining and Knowledge Discovery*, vol. 28, no. 4, pp. 851–881, 2014.

[148] L. Ye and E. Keogh, "Time series shapelets: A novel technique that allows accurate, interpretable and fast classification," *Data Mining and Knowledge Discovery*, vol. 22, no. 1, pp. 149–182, 2011.

[149] W. Zhang and others, "Time series classification by shapelet dictionary learning with SVM-based ensemble classifier," *Computational Intelligence and Neuroscience*, vol. 2021, p. 5586273, 2021.

[150] J. M. Mendel and R. B. John, “Type-2 fuzzy sets made simple,” *IEEE Transactions on fuzzy systems*, vol. 10, no. 2, pp. 117–127, 2002, publisher: IEEE. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/995115/

[151] I. Škrjanc, J. A. Iglesias, A. Sanchis, D. Leite, E. Lughofer, and F. Gomide, “Evolving cloud-based system for the recognition of drivers’ stress and drowsiness based on a fuzzy logic system,” *Engineering Applications of Artificial Intelligence*, vol. 168, pp. 448–464, 2019.

[152] F. J. Baldán and L. Martínez, “Time series features and fuzzy memberships combination for time series classification,” *Neurocomputing*, vol. 606, p. 128368, Nov. 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925231224011391

[153] S. Chakraborty and F. Heintz, “Enhancing Time Series Forecasting with Fuzzy Attention-Integrated Transformers,” Mar. 2025, arXiv:2504.00070 [cs]. [Online]. Available: http://arxiv.org/abs/2504.00070

[154] J. Lines, S. Taylor, and A. Bagnall, “Time series classification with HIVE-COTE: The hierarchical vote collective of transformation-based ensembles,” *ACM Transactions on Knowledge Discovery from Data*, vol. 12, no. 5, pp. 1–35, 2018.

[155] M. Middlehurst, J. Large, M. Flynn, J. Lines, A. Bostrom, and A. Bagnall, “HIVE-COTE 2.0: a new meta ensemble for time series classification,” *Machine Learning*, vol. 110, no. 11, pp. 3211–3243, 2021.

[156] C. W. Tan, A. Dempster, C. Bergmeir, and G. I. Webb, “MultiROCKET: Multiple pooling operators and transformations for fast and effective time series classification,” *Data Mining and Knowledge Discovery*, vol. 36, no. 5, pp. 1623–1646, 2022.

[157] N. Fronda and H. Abbas, “Differentiable inference of temporal logic formulas,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 41, no. 11, pp. 4193–4204, 2022.

[158] D. Li, M. Cai, C.-I. Vasile, and R. Tron, “TLINet: Differentiable neural network temporal logic inference,” *arXiv preprint arXiv:2405.06670*, 2024.

[159] S. Guillaume and B. Charnomordic, “Learning interpretable fuzzy inference systems with FisPro,” *Information Sciences*, vol. 181, pp. 4409–4427, 2011.

[160] C. Chang, W.-C. Peng, and T.-F. Chen, “Llm4ts: Two-stage fine-tuning for time-series forecasting with pre-trained llms,” *CoRR*, 2023. [Online]. Available: https://openreview.net/forum?id=6MKvV3bpfk

[161] J. D. Hamilton, *Time series analysis*. Princeton university press, 2020. [Online]. Available: https://books.google.com/books?hl=en&lr=&id=BeryDwAAQBAJ&oi=fnd&pg=PP1&dq=hamilton+time+series+analysis+&ots=BhCV704Udm&sig=A3Sh1THv2zSjwsdz43O04ccuFHY

[162] A. J. Fox, “Outliers in time series,” *Journal of the Royal Statistical Society Series B: Statistical Methodology*, vol. 34, no. 3, pp. 350–363, 1972, publisher: Oxford University Press. [Online]. Available: https://academic.oup.com/jrsssb/article-abstract/34/3/350/7027249

[163] “Time Series Classification Website,” 2023. [Online]. Available: https://timeseriesclassification.com/

[164] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley, 1989.

[165] M. Middlehurst, “TSML-EVAL.” [Online]. Available: https://github.com/time-series-machine-learning/tsml-eval

[166] X. Zhang, Y. Gao, J. Lin, and C.-T. Lu, “TapNet: Multivariate time series classification with attentional prototypical network,” in *Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence*, vol. 34, 2020, pp. 6845–6852, issue: 4.

[167] J. Morrill, A. Fermanian, P. Kidger, and T. Lyons, “A Generalised Signature Method for Multivariate Time Series Feature Extraction,” Feb. 2021, arXiv:2006.00873 [cs]. [Online]. Available: http://arxiv.org/abs/2006.00873

[168] A. Bhatia and H. Hagras, “A time series based explainable interval type-2 fuzzy logic system,” in *2022 IEEE International Conference on Fuzzy*

*Systems (FUZZ-IEEE).* IEEE, 2022, pp. 1–7. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9882556/

[169] ——, "A Type-2 Fuzzy Time Series Classification System with Optimized Time Period Selection," in *2024 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE).* IEEE, 2024, pp. 1–9. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10611905/

[170] S. Usmankhujaev, B. Ibrokhimov, S. Baydadaev, and J. Kwon, "Time series classification with inceptionfcn," *Sensors*, vol. 22, no. 1, 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/1/157

[171] Z. Huang and Q. Shen, "Fuzzy interpolative reasoning via scale and move transformations," *IEEE Transactions on Fuzzy Systems*, vol. 16, no. 2, pp. 340–359, 2006.

[172] ——, "Fuzzy interpolative and extrapolative reasoning: A practical approach," *IEEE Transactions on Fuzzy Systems*, vol. 16, no. 1, pp. 13–28, 2008.

[173] F. Li, C. Shang, Y. Li, J. Yang, and Q. Shen, "Approximate reasoning with fuzzy rule interpolation: Background and recent advances," *Artificial Intelligence Review*, vol. 54, pp. 4543–4590, 2021.

[174] ——, "Interpolation with just two nearest neighbouring weighted fuzzy rules," *IEEE Transactions on Fuzzy Systems*, vol. 28, no. 9, pp. 2255–2262, 2020.

[175] F. Li, Y. Li, C. Shang, and Q. Shen, "Fuzzy knowledge-based prediction through weighted rule interpolation," *IEEE Transactions on Cybernetics*, vol. 50, no. 10, pp. 4508–4517, 2020.

[176] P. Zhang, C. Shang, and Q. Shen, "Fuzzy rule interpolation with k-neighbours for tsk models," *IEEE Transactions on Fuzzy Systems*, vol. 30, no. 10, pp. 4031–4043, 2022.

[177] C. Jiang, C. Shang, and Q. Shen, "Reinforcing tsk model interpolation with rule weight adjustment via location view analysis," *IEEE Transactions on Fuzzy Systems*, vol. 33, no. 2, pp. 580–592, 2025.

[178] M. Zhou, C. Shang, and Q. Shen, "Rule-ranking-based approximate knowledge interpolation with directional monotonicity," *IEEE Transactions on Cybernetics*, vol. 54, no. 8, pp. 4814–4827, 2024.

[179] J. Yang, C. Shang, Y. Li, F. Li, and Q. Shen, "Anfis construction with sparse data via group rule interpolation," *IEEE Transactions on Cybernetics*, vol. 51, no. 5, pp. 2773–2786, 2021.

[180] J. Yang, C. Shang, Y. Li, F. Li, L. Shen, and Q. Shen, "Constructing anfis with sparse data through group-based rule interpolation: An evolutionary approach," *IEEE Transactions on Fuzzy Systems*, vol. 30, no. 4, pp. 893–907, 2022.

[181] N. Chukhrova and A. Johannssen, "Fuzzy Regression Analysis: Systematic Review and Bibliography," *Applied Soft Computing*, vol. 84, p. 105708, 2019.

[182] A.-T. Nguyen, T. Taniguchi, L. Eciolaza, V. Campos, R. Palhares, and M. Sugeno, "Fuzzy Control Systems: Past, Present and Future," *IEEE Computational Intelligence Magazine*, vol. 14, no. 1, pp. 56–68, 2019.

[183] wordsforthewise, "All Lending Club Loan Data," 2019, published: Kaggle Dataset. https://www.kaggle.com/datasets/wordsforthewise/lending-club.

[184] R. Hamon, H. Junklewitz, and I. Sanchez, "Robustness and Explainability of Artificial Intelligence: From Technical to Policy Solutions," *JRC Technical Report*, 2020, publisher: European Commission.

[185] X. Gu, P. Angelov, and Q. Shen, "Semi-supervised fuzzily weighted adaptive boosting for classification," *IEEE Transactions on Fuzzy Systems*, vol. 32, no. 4, pp. 2318–2330, 2024.

[186] H. Lv, F. Li, C. Shang, and Q. Shen, "W-infer-polation: Approximate reasoning via integrating weighted fuzzy rule inference and interpolation," *Knowledge-Based Systems*, vol. 258, p. 109995, 2022.

[187] J. Marín-Blázquez and Q. Shen, "From approximative to descriptive fuzzy classifiers," *IEEE Transactions on Fuzzy Systems*, vol. 10, no. 4, pp. 484–497, 2002.

[188] J. Sanz, D. Bernardo, F. Herrera, H. Bustince, and H. Hagras, "A compact evolutionary interval-valued fuzzy rule-based classification system for the modeling and prediction of real-world financial applications with imbalanced data," *IEEE Transactions on Fuzzy Systems*, vol. 23, pp. 973–990, 2015.

[189] M. Dorigo, V. Maniezzo, and A. Colorni, "Ant system: Optimization by a colony of cooperating agents," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 26, pp. 29–41, 1996.

[190] Y. Chen, "Study on centroid type-reduction of interval type-2 fuzzy logic systems based on noniterative algorithms," *Complexity*, vol. 2019, pp. 1–12, 2019.

[191] J. Alcalá-Fdez, L. Sánchez, S. García, M. del Jesus, S. Ventura, J. Garrell, J. Otero, C. Romero, J. Bacardit, V. Rivas, J. Fernández, and F. Herrera, "KEEL: A software tool to assess evolutionary algorithms for data mining problems," *Journal of Multiple-Valued Logic and Soft Computing*, vol. 17, pp. 255–287, 2011.

[192] R. Alcalá, J. Alcalá-Fdez, F. Herrera, and J. Otero, "Genetic learning of accurate and compact fuzzy rule based systems based on the 2-tuples linguistic representation," *International Journal of Approximate Reasoning*, vol. 44, pp. 45–64, 2007.

[193] M. Dorigo and T. Stützle, *Ant Colony Optimization*. Cambridge, MA: MIT Press, 2004.

[194] M. Dorigo, M. Birattari, and T. Stützle, "Ant colony optimization," *IEEE Computational Intelligence Magazine*, vol. 1, no. 4, pp. 28–39, 2006.

[195] R. Jensen and Q. Shen, "Finding rough set reducts with ant colony optimization," in *Proceedings of the 2003 UK Workshop on Computational Intelligence*, 2003, pp. 15–22.

[196] T. Stützle and H. H. Hoos, "MAX–MIN ant system," *Future Generation Computer Systems*, vol. 16, no. 8, pp. 889–914, 2000.

[197] G. Vilone and L. Longo, "A quantitative evaluation of global, rule-based explanations of post-hoc, model agnostic methods," *Frontiers in Artificial Intelligence*, vol. 4, p. 717899, 2021.