Autonomous learning for fuzzy systems: a review

Xiaowei Gu[1] · Jungong Han[2] · Qiang Shen[2] · Plamen P. Angelov[3]

Abstract

As one of the three pillars in computational intelligence, fuzzy systems are a powerful mathematical tool widely used for modelling nonlinear problems with uncertainties. Fuzzy systems take the form of linguistic IF-THEN fuzzy rules that are easy to understand for human. In this sense, fuzzy inference mechanisms have been developed to mimic human reasoning and decision-making. From a data analytic perspective, fuzzy systems provide an effective solution to build precise predictive models from imprecise data with great transparency and interpretability, thus facilitating a wide range of real-world applications. This paper presents a systematic review of modern methods for autonomously learning fuzzy systems from data, with an emphasis on the structure and parameter learning schemes of mainstream evolving, evolutionary, reinforcement learning-based fuzzy systems. The main purpose of this paper is to introduce the underlying concepts, underpinning methodologies, as well as outstanding performances of the state-of-the-art methods. It serves as a one-stop guide for readers learning the representative methodologies and foundations of fuzzy systems or who desire to apply fuzzy-based autonomous learning in other scientific disciplines and applied fields.

Keywords Fuzzy systems · Autonomous learning · Evolving · Evolutionary · Reinforcement learning

✉ Jungong Han
  juh22@aber.ac.uk

  Xiaowei Gu
  x.gu@kent.ac.uk

  Qiang Shen
  qqs@aber.ac.uk

  Plamen P. Angelov
  p.angelov@lancaster.ac.uk

[1] School of Computing, University of Kent, Canterbury CT2 7NZ, UK

[2] Department of Computer Science, Aberystwyth University, Aberystwyth SY23 3DB, UK

[3] School of Computing and Communications, Lancaster University, Lancaster LA1 4WA, UK

1 Introduction

In recent years, deep neural networks (DNNs) have gained huge popularity among researchers and the general public thanks to the eye-catching performances they have demonstrated on many challenging tasks involving visual and audio information (LeCun et al. 2015; Litjens et al. 2017; Li et al. 2018). The breakthroughs achieved by DNNs have also led to the recent hot wave of artificial intelligence (AI) and machine learning research. With the exponentially increasing amount of data for analysis, the stronger capability and greater availability of computational hardware, powerful AI models have been constructed, clearly illustrating super-human machine intelligence as the impetus to revolutionize our industries, society, and everyday lives (Angelov and Gu 2018a; Hagras 2018).

Despite the great promise for a technology revolution, concerns on the issues of *understandability* and *trustability* have been raised frequently in a wide range of research communities and industries from the use of complicated machine learning systems (Angelov and Gu 2018a; Hagras 2018; Angelov et al. 2021). For example, the state-of-the-art DNNs are often characterized as the typical type of "black box" models that lack transparency (Zhou and Feng 2017). Such models usually have a huge number of hyper-parameters learned from very large amounts of training data but with no clear physical meaning. Their learning performances rely on careful parameter tuning whilst the training and internal reasoning processes are hardly understandable nor interpretable to human (Feng et al. 2018). The "black box" nature is not specific to DNNs, but also to many mainstream models such as support vector machine (Cristianini and Shawe-Taylor 2000), random forest (Breiman 2001), and learning vector quantization (Kohonen and Maps 1995). Whilst decision tree (Quinlan 1986) and k-nearest neighbour (Cunningham and Delany 2007) are known to be interpretable on small-scale, simple problems, their interpretability and explainability are usually very limited when applied for large-scale, high-dimensional, complex problems (Hagras 2018). Explaining the decisions made by these state-of-the-art models is often challenging and problematic. The lack of explainable features has also hindered their applicability in financial, healthcare and other safety–critical applications, where being able to understand the rationale behind the models' decision is the key for trust (Dosilovic et al. 2018; Tjoa and Guan 2020).

Fuzzy rule-based systems provide an effective solution for constructing models that offer both great prediction precision and high-level model transparency while being capable of incorporating real-world uncertainties (Angelov and Gu 2018a; Barredo Arrieta, et al. 2020). At the core of such systems is the fuzzy set theory defined by Lofti A. Zadeh in the mid-60 s (Zadeh 1965). Built upon solid theoretical foundation, fuzzy rule-based systems attempt to mimic human reasoning and decision-making processes rather than trying to represent the human brain (Hagras 2018). The use of linguistic IF-THEN fuzzy rules in fuzzy rule-based systems is a natural and intuitive representation of knowledge easy-to-interpret by human. The human-like fuzzy inferencing process enables an explicit explanation for each decision made by fuzzy rule-based models. The decisions can be examined and verified by domain experts, and direct human interaction is also facilitated. With a transparent system structure, an explainable internal reasoning mechanism and intuitive IF-THEN rules to represent abstract knowledge, fuzzy rule-based systems have increasingly regained attention in the current move towards *explainable AI* (Garibaldi 2019).

This paper provides a systematic review of modern methods for autonomously constructing zero-order and first-order fuzzy rule-based predictive models from data. This review is mostly focused on the autonomous learning features of three different classes of

fuzzy rule-based systems, including evolving fuzzy systems, evolutionary fuzzy systems, and reinforcement learning-based fuzzy systems. The main purpose of this overview is to recall the basics of fuzzy systems, to revisit and evaluate different approaches for building such systems, and to summarize the related applications tackling real-world problems. The latest trends and challenges for the development of fuzzy models will also be analysed, which may be useful to learners intended to quickly grasp the current progress in this exciting research area as a whole. Due to the wider range of topics covered, this paper does not aim to provide an exhaustive list of related works but instead focuses on a smaller portion of more representative and popular ones, in order to present the general concepts and principles. Interested readers are referred to previous reviews (Herrera 2008; Fazzolari et al. 2013; Fernández et al. 2015; Fernandez et al. 2019; Škrjanc et al. 2019; Campos Souza 2020; Leite et al. 2020) and monographs (Cordón et al. 2001a; Kasabov 2007; Angelov et al. 2010; Lughofer 2011a; Angelov 2012; Angelov and Gu 2018b) for more details about the developments of fuzzy systems and their applications.

To summarize, main contributions of this paper are as follows:

1. An overview of the basic concepts of fuzzy sets and zero-order and first-order fuzzy systems.
2. A systematic review of the autonomous learning of evolving, evolutionary and reinforcement learning-based fuzzy systems.
3. A critical comparison amongst the three classes of fuzzy systems from learning and application perspectives.
4. A comprehensive discussion on the open issues of fuzzy systems research, highlighting directions for future development.

The remainder of this paper is organized as follows. Section 2 recalls the basic concepts of fuzzy sets and fuzzy systems. The structure and parameter learning schemes of mainstream evolving fuzzy systems are reviewed in Sect. 3. Section 4 summarizes the evolutionary tuning and learning schemes employed by mainstream evolutionary fuzzy systems. Section 5 describes the use of reinforcement learning in constructing fuzzy systems. A comparison between the aforementioned three classes of fuzzy systems is presented in Sect. 6. Section 7 introduces the real-world applications where these fuzzy systems have been implemented for. A case study on stock price prediction is presented in Sect. 8. In the end, this paper is concluded with a discussion about challenges and future directions.

2 Preliminaries: fuzzy sets and fuzzy systems

In this section, basic concepts of fuzzy sets and fuzzy rule-based systems are presented.

2.1 Fuzzy sets

Fuzzy set theory was introduced by (Zadeh 1965) to generalize the concept of fuzziness that is contained in human language, judgment, evaluation, and decisions naturally. Fuzzy sets extend classical set theory, where a data sample can partially belong to multiple sets at the same time with different degrees of membership. Compared with classical bivalent sets (or crisp sets), fuzzy sets allow an extra degree of flexibility to incorporate real-world uncertainties (Zimmermann 2010).
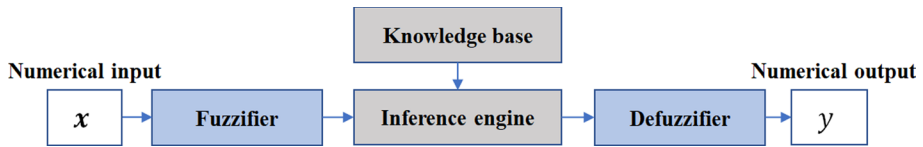
Fig. 1 Standard structure of fuzzy rule-based systems

A fuzzy set, *A* is characterized by a membership function denoted by $\mu_A(x)$ with a value range of [0, 1] (Zadeh 1965). The value of $\mu_A(x)$ represents the degree of truth, fulfilment, satisfaction of the fact that *x* belongs to *A* (Angelov and Gu 2018b). A higher value of $\mu_A(x)$, namely, closer to 1, means that the data sample *x* is more likely to be a full member of *A*, and vice versa. In the extreme cases where $\mu_A(x) = 1$ or 0, which means *x* is or is not belonging to *A*, the fuzzy set reduces to the crisp set with a Boolean membership function. Therefore, fuzzy sets can be viewed as a more generalized form of classical bivalent sets.

The commonly used membership functions in fuzzy sets include, but are not limited to, Gaussian, triangular, trapezoidal, Cauchy, etc. Among these, the Gaussian type membership function, as formulated by Eq. (1), is typically considered to be the most widely used one, thanks to the good generalization capabilities and coverage of the whole feature space (Angelov and Zhou 2008).

$$\mu_A(x) = e^{-\frac{(x-p)^2}{2\sigma^2}} \tag{1}$$

where *p* and *σ* are the respective focal point and spread of the fuzzy set *A*.

In practice, the type of membership function to be used is usually determined by human experts based on their preferences and experience, and may differ significantly from real data distribution (Angelov and Yager 2012). Parameters of membership functions can be either handcrafted with human expertise or learned from data by learning algorithms (Angelov and Gu 2018b).

2.2 Fuzzy systems

Fuzzy rule-based systems are one of the most important application fields of fuzzy set theory. Fuzzy rule-based systems can be interpreted as a set of intuitive and easy-to-interpret IF-THEN fuzzy rules. Fuzzy rule-based systems can also be represented as neuro-fuzzy systems, which can be viewed as a special type of multi-layer feedforward neural networks.

A standard fuzzy rule-based system is composed of four functional components, namely, fuzzifier (maps numerical crisp inputs to fuzzy inputs), inference engine (maps fuzzy inputs to fuzzy outputs using fuzzy rules), knowledge base (contains all fuzzy rules) and defuzzifier (maps fuzzy outputs to numerical crisp outputs) (Fernandez et al. 2019), see Fig. 1.

The knowledge base is the key component for decision-making, and it is represented by intuitive and easy-to-interpret IF-THEN fuzzy rules. Fuzzy rules are usually composed of two parts, namely, premise (IF) part stating conditions on the input(s) and consequent (THEN) part describing the corresponding output(s) (Adriaenssens et al. 2004).

There are two most widely used types of fuzzy rules:

(1) Zadeh-Mamdani type (Zadeh 1973; Mamdani and Assilian 1975), introduced by Zadeh, Mamdani and Assilian around 1970s.

(2) Takagi-Sugeno-Kang type (Takagi and Sugeno 1985; Sugeno and Kang 1988), introduced by Takagi, Sugeno and Kang around 1980s.

A typical Zadeh-Mamdani type fuzzy rule-based system is composed of fuzzy rules in the form of Eq. (2):

$$\mathbf{R}_i : \quad \begin{array}{c} IF \ \left(x_1 \ is \ A_{i,1}\right) AND \ \left(x_2 \ is \ A_{i,2}\right) AND \ \dots \ AND \ \left(x_N \ is \ A_{i,N}\right) \\ THEN \ \left(y_i \ is \ A_{i,out}\right) \end{array} \tag{2}$$

where $\mathbf{R}_i$ is the ith fuzzy rule; $\boldsymbol{x} = [x_1, x_2, \dots, x_N]^T$ is the $N \times 1$ dimensional system input; x_n is the nth input variable; $A_{i,n}$ is the linguistic term of the nth premise fuzzy set of $\mathbf{R}_i$; $A_{i,out}$ is the linguistic term of the output; and y_i is the output of $\mathbf{R}_i$.

In contrast, a typical Takagi-Sugeno-Kang type fuzzy rule-based system consists of fuzzy rules in the following form:

$$\mathbf{R}_i : \quad \begin{array}{c} IF \ \left(x_1 \ is \ A_{i,1}\right) AND \ \left(x_2 \ is \ A_{i,2}\right) AND \ \dots \ AND \ \left(x_N \ is \ A_{i,N}\right) \\ THEN \ \left(y_i = \boldsymbol{a}_i^T \bar{\boldsymbol{x}}\right) \end{array} \tag{3}$$

where $\bar{\boldsymbol{x}} = [1, \boldsymbol{x}^T]^T$; $\boldsymbol{a}_i = [a_0, a_1, a_2 \dots, a_N]^T$ is the $(N+1) \times 1$ dimensional consequent parameter vector of $\mathbf{R}_i$.

It is clear from Eqs. (2) and (3) that both Zadeh-Mamdani type and Takagi-Sugeno-Kang type fuzzy rules have the same linguistic forms of premise parts, but differ in the consequent parts. The consequent part of a Takagi-Sugeno-Kang type fuzzy rule is a linear regression function, while the consequent part of a Zadeh-Mamdani type fuzzy rule is a singleton. Thus, Zadeh-Mamdani type fuzzy rules are also called zero-order fuzzy rules, while Takagi–Sugeno-Kang type fuzzy rules are called first-order fuzzy rules.

In (Angelov and Yager 2012), a novel type of fuzzy rules called AnYa was proposed to simplify the antecedent parts of fuzzy rules to numerical vectors as prototypes learned from data. This type of fuzzy rules also removes the need of defining membership functions for each variable. A zero-order AnYa type fuzzy rule can be formulated as follows (Angelov and Yager 2012),

$$\mathbf{R}_i : \quad \begin{array}{c} IF\left(\boldsymbol{x} \sim \boldsymbol{p}_i\right) \\ THEN\left(y_i \ is \ A_{i,out}\right) \end{array} \tag{4}$$

and a first-order AnYa type fuzzy rule is expressed as:

$$\mathbf{R}_i : \quad \begin{array}{c} IF\left(\boldsymbol{x} \sim \boldsymbol{p}_i\right) \\ THEN\left(y_i = \boldsymbol{a}_i^T \bar{\boldsymbol{x}}\right) \end{array} \tag{5}$$

where $\boldsymbol{p}_i$ is the prototype of $\mathbf{R}_i$; "$\sim$" denotes similarity. One can see from Eqs. (4) and (5) that AnYa type fuzzy rules have the same consequent parts as the Zadeh-Mamdani type and Takagi–Sugeno-Kang type fuzzy rules. AnYa fuzzy rule-based systems also follow the same standard fuzzy inference procedure.

2.3 Fuzzy systems for prediction and classification

A typical fuzzy predictor is a system that uses first-order fuzzy rules. Taken the Takagi-Sugeno-Kang type fuzzy rules (Eq. (3)) as an example, the input–output relationship of a

standard multi-input single-output fuzzy system is mathematically modelled as (Angelov 2012; Angelov and Gu 2018b):

$$\widehat{y} = f(\boldsymbol{x}) = \sum_{i=1}^{L} \overline{\lambda}_i y_i = \sum_{i=1}^{L} \overline{\lambda}_i \boldsymbol{a}_i^T \overline{\boldsymbol{x}} \tag{6}$$

where L is the number of fuzzy rules in the system; y_i is the output of the ith fuzzy rule, $\mathbf{R}_i$; $\overline{\lambda}_i$ is the normalized firing strength of $\mathbf{R}_i$ calculated by Eq. (7):

$$\overline{\lambda}_i = \frac{\lambda_i}{\sum_{j=1}^{L} \lambda_j} \tag{7}$$

It can be seen from Eqs. (6) and (7) that given a particular input, $\boldsymbol{x}$, the overall system output is computed as a fuzzily weighted sum of individual fuzzy rules' outputs following the so-called "centre of gravity" principle. The fuzzy weights are defined by the firing strengths of the fuzzy rules (Angelov 2012; Angelov and Yager 2012; Angelov and Gu 2018b).

The most commonly used way to calculate the firing strength, λ_i of a fuzzy rule $\mathbf{R}_i$ based on the membership values is the t-norm:

$$\lambda_i = \prod_{n=1}^{N} \mu_{A_{i,n}}(x_n) \tag{8}$$

Note that, t-norm is dominantly utilized by fuzzy systems with axis parallel rules to aggregate membership values of different attributes (Angelov and Filev 2004; Márquez et al. 2007). For fuzzy systems with non-axis parallel rules, multivariate Gaussian type membership functions are typically employed, and the membership values produced by multivariate membership functions are used directly as the firing strengths of the fuzzy rules (Lemos et al. 2011; Lughofer et al. 2015; Pratama et al. 2015).

A multi-input single-output (MISO) first-order fuzzy predictor can be directly applied to binary classification problems by using the following simple rule to convert continuous system outputs into discrete class labels ("0" and "1"):

$$\widehat{y} = \begin{cases} \text{class } 1 & \text{if } f(\boldsymbol{x}) > 0.5 \\ \text{class } 0 & \text{else} \end{cases} \tag{9}$$

A MISO fuzzy predictor can also be generalized to a multi-input multi-output (MIMO) one for multi-class classification problems:

$$\widehat{\boldsymbol{y}} = f(\boldsymbol{x}) = \sum_{i=1}^{L} \overline{\lambda}_i \boldsymbol{y}_i = \sum_{i=1}^{L} \overline{\lambda}_i \mathbf{a}_i^T \overline{\boldsymbol{x}} \tag{10}$$

where $\widehat{\boldsymbol{y}} = \left[\widehat{y}_1, \widehat{y}_2, \ldots, \widehat{y}_W\right]^T$ is the $W \times 1$ dimensional system output; $\mathbf{a}_i = [\boldsymbol{a}_{i,1}, \boldsymbol{a}_{i,2}, \ldots, \boldsymbol{a}_{i,W}]$ is a $(N+1) \times W$ dimensional consequent parameter matrix of $\mathbf{R}_i$; $\boldsymbol{y}_i$ is the $W \times 1$ dimensional output of $\mathbf{R}_i$; and the MIMO Takagi-Sugeno-Kang type fuzzy rule, $\mathbf{R}_i$ takes the following form (Angelov 2010):

$$\mathbf{R}_i: \begin{array}{c} IF \left(x_1 \text{ is } A_{i,1}\right) AND \left(x_2 \text{ is } A_{i,2}\right) AND \ldots AND \left(x_N \text{ is } A_{i,N}\right) \\ THEN \left(y_{i,1} = \boldsymbol{a}_{i,1}^T \overline{\boldsymbol{x}}\right) AND \left(y_{i,2} = \boldsymbol{a}_{i,2}^T \overline{\boldsymbol{x}}\right) AND \ldots AND \left(y_{i,W} = \boldsymbol{a}_{i,W}^T \overline{\boldsymbol{x}}\right) \end{array} \tag{11}$$

Accordingly, the class labels are determined by:

$$\widehat{y} = \text{class } i^*; \ i^* = \underset{i=1,2,\ldots,W}{\text{argmax}} \left(\widehat{y}_i \right) \tag{12}$$

Comparing with first-order fuzzy rules, zero-order fuzzy rules Eq. (2) are more widely used for constructing classification models. For a standard zero-order fuzzy rule-based classifier, the class label of a given input sample, x is determined by the fuzzy rule that produces the highest firing strength calculated by Eq. (8) (Angelov and Gu 2018b; Angelov and Zhou 2008; Ishibuchi et al. 1995):

$$\widehat{y} = A_{i^*,out}; \ i^* = \underset{i=1,2,\ldots,L}{\text{argmax}} \left(\lambda_i \right) \tag{13}$$

where the linguistic terms of zero-order fuzzy rules, $A_{1,out}$, $A_{2,out}$,…,$A_{L,out}$ are the class labels.

3 Evolving fuzzy systems

The concept of evolving fuzzy systems was firstly conceived around the beginning of twenty-first century (Angelov and Buswell 2001, 2002; Kasabov and Song 2002; Angelov 2002; Angelov and Filev 2002, 2003). Evolving fuzzy systems are a class of fuzzy systems that are capable of self-developing and self-updating the system structure and parameters online from data streams (Ge and Zeng 2020; Gu and Shen 2021). A typical evolving fuzzy system can learn from streaming data "on the fly" in a single-pass manner, efficiently transforming the learned knowledge into human-interpretable fuzzy rules. It is capable of capturing concept drifts and/or shifts in the data streams and self-evolving its structure and parameters to self-adapt to the dynamically changing data patterns (Lughofer and Angelov 2011). As an effective and promising tool for handling streaming data problems, evolving fuzzy systems have been intensively researched in the past two decades.

The general framework of evolving fuzzy system is depicted in Fig. 2, where one can see that the evolving mechanism of a typical evolving fuzzy system consists of the following two key schemes, namely (1) structure evolving, and (2) parameter updating (Ge and Zeng 2018a, 2020; Rong et al. 2018). The structure evolving scheme mostly concerns fuzzy rule generation, merging, pruning and splitting as well as the premise parameter learning, namely prototypes of fuzzy rules. The parameter learning scheme is mostly for learning the meta-parameters of the system and consequent parameters of fuzzy rules. In the rest of this section, a review of evolving mechanisms used by a selected group of highly representative evolving fuzzy systems in the literature is presented. An index of the evolving fuzzy systems reviewed in this section is presented in Table 1 for readers' convenience. Since a large part of the literature has been reviewed recently by Škrjanc et al. (2019); Campos Souza 2020; Leite et al. 2020), this section is more focused on general concepts and principles. As stated in Sect. 1, this section only considers zero-order and first-order evolving fuzzy systems. For higher order ones, such as evolving type-2 fuzzy systems, more details can be found in Škrjanc et al. (2019).
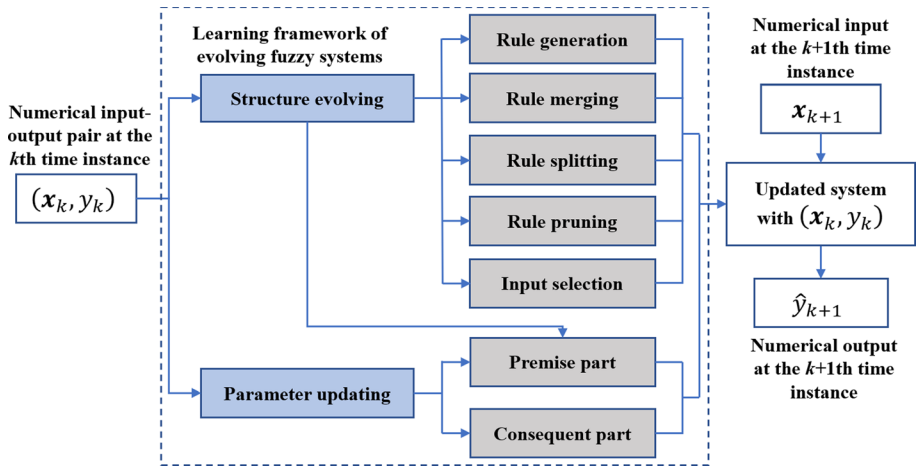
Fig. 2 General framework of evolving fuzzy systems (Ge and Zeng 2020)

3.1 Structure learning schemes

Evolving fuzzy systems may employ different structure evolving schemes for fuzzy rule identification. The key idea of such structure evolving schemes is to group streaming data into clusters using recursive clustering and associate each individual cluster with a particular fuzzy rule in the rule base. In this way, the identified local models of data are converted into human-interpretable fuzzy rules.

3.1.1 Rule generation

Rule generation is the key component of the structure evolving scheme because evolving fuzzy systems begin with no fuzzy rule in the knowledge base. The fuzzy rules are added to the knowledge base one-by-one during the learning process from streaming data "on the fly" in an exploratory way to capture the concept drifts and shifts (Lughofer and Angelov 2011). Typically, an evolving fuzzy system adds new fuzzy rules when unfamiliar data patterns are observed from new data samples. As the existing fuzzy rules in the knowledge base fail to describe the novel data patterns well, the system needs to construct new fuzzy rules to self-adapt to the changes. Different evolving fuzzy systems use different approaches to identify new fuzzy rules from data streams (Rong et al. 2018). To date, the most widely used approaches include density/potential criterion, distance criterion, error criterion, firing strength criterion and statistical contribution criterion (Rong et al. 2018; Bao et al. 2018). In this subsection, the listed evolving fuzzy systems in Table 1 will be categorized according to the criteria they utilize for fuzzy rule generation.

3.1.1.1 Density/potential criterion Evolving Takagi-Sugeno (eTS) model (Angelov and Filev 2004), as one of the earliest evolving fuzzy systems, employs the potential criterion to identify new fuzzy rules. The potential of a data sample is calculated based on the average Euclidean distance between this sample and all other samples in the data space in the form of Cauchy function. If the potential of a new data sample is greater than all the existing clus-

Table 1 Index of evolving fuzzy systems covered in this review

Model	Abbreviation	Year
Evolving fuzzy neural network (Kasabov 2001)	EFuNN	2001
Evolving Takagi-Sugeno fuzzy model (Angelov and Filev 2002, 2003, 2004; Angelov and Buswell 2001, 2002)	eTS	
Self-constructing fuzzy neural network (Lin et al. 2001)	SCFNN	
Dynamic evolving neural-fuzzy inference system (Kasabov and Song 2002)	DENFIS	2002
Self-organizing fuzzy neural networks (Leng et al. 2004, 2005)	SOFNN	2004
Simplified evolving Takagi-Sugeno fuzzy model (Angelov and Filev 2005)	Simpl_eTS	2005
Sequential adaptive fuzzy inference system (Rong et al. 2006)	SAFIS	2006
Zero-order evolving Takagi-Sugeno fuzzy classifier (Angelov and Zhou 2008)	eClass0	2008
First-order evolving Takagi-Sugeno fuzzy classifier (Angelov and Zhou 2008)	eClass1	
Flexible fuzzy inference system (Lughofer 2008)	FLEXFIS	
Self-organizing fuzzy modified least-squares network (Jesús Rubio 2009)	SOFMLS	2009
Evolving neuro-fuzzy model (Soleimani-B et al. 2010)	ENFM	2010
Gaussian evolving fuzzy modelling system (Lemos et al. 2011)	eMG	2011
Extended sequential adaptive fuzzy inference system (Rong et al. 2011)	ESAFIS	
Simplified evolving Takagi-Sugeno neural-fuzzy system (Angelov 2011)	Simpl_eTS+	
Meta-cognitive neuro-fuzzy inference system (Subramanian and Suresh 2012)	McFIS	2012
Generic evolving neuro-fuzzy inference system (Pratama et al. 2014a)	GENEFIS	2014
Parsimonious network based on fuzzy inference system (Pratama et al. 2014b)	PANFIS	
Evolving fuzzy model (Dovžan et al. 2015)	EFuMo	2015
Generalized smart evolving fuzzy system (Lughofer et al. 2015)	GS-EFS	
Generic self-evolving Takagi–Sugeno-Kang fuzzy neural network (Nguyen et al. 2015)	GSETSK	
Evolving possibilistic fuzzy modelling system (Maciel et al. 2017)	ePFM	2017
Autonomous learning multi-model system (Angelov et al. 2018)	ALMMo	2018
Correntropy-based evolving fuzzy neural system (Bao et al. 2018)	CEFNS	
Local error optimization approach for learning evolving fuzzy system (Ge and Zeng 2018a)	LEOA	
Self-organizing fuzzy inference system (Gu and Angelov 2018a)	SOFIS	
Parsimonious learning machine (Ferdaus et al. 2019)	PALM	2019
Recursive maximum correntropy-based evolving fuzzy system (Rong et al. 2019)	RMCEFS	
Self-evolving fuzzy system (Ge and Zeng 2018b)	SEFS	
Spatio-temporal fuzzy inference system (Samanta et al. 2019)	SPATFIS	
Evolving fuzzy system with self-learning/adaptive thresholds (Ge and Zeng 2020)	EFS-SLAT	2020
Incremental fuzzy c-regression clustering-based system (Blazic and Skrjanc 2020)	InFuR	
Jointly evolving and compressing fuzzy system (Huang et al. 2021)	JECFS	2021
Statistically evolving fuzzy inference system (Yang et al. 2022a)	SEFIS	

ter centres, a new cluster is identified with this new sample as its focal point, namely, cluster centre, and a new fuzzy rule is initialized accordingly. The evolving fuzzy rule-based classifiers, eClass1 and eClass0 (Angelov and Zhou 2008) also employ the potential criterion for fuzzy rule identification in the same way as eTS. The simplified evolving Takagi-Sugeno (Simpl_eTS) model (Angelov and Filev 2005) uses a simplified variation of potential named scatter. The scatter of a data sample is defined as the average Euclidean distance between

this sample and all other samples. A new data sample is recognized as the focal point of a new cluster if its scatter value is greater or smaller than all existing focal points. Meta-cognitive neuro-fuzzy inference system (McFIS) (Subramanian and Suresh 2012) use both the potential and error criteria for new rule identification. If the prediction error on the current input sample and its spherical potential, which is defined as the absolute value of the quantity in squared distance mapping, are both greater than the predefined thresholds, a new rule will be initialized. In the autonomous learning multi-model (ALMMo) system (Angelov et al. 2018), another variation of potential called global density is used for fuzzy rule identification. The global density is calculated based on the Euclidean distance between the data sample and the arithmetic mean of all the observed data samples in the entire data space. The global density measures the similarity between a data sample and the global data pattern. A new fuzzy rule will be initialized by a new data sample if its global density is smaller or greater than existing focal points. If this condition is satisfied, it suggests that the new sample describes the global data pattern better than others (in this case, its global density is currently the greatest) or it represents a new unfamiliar pattern that cannot be represented by existing focal points (in this case, its global density is currently the smallest). Self-organizing fuzzy inference system (SOFIS) (Gu and Angelov 2018a) initializes new fuzzy rules by identifying identifies the focal points from data based on their multimodal density values and mutual distances. The multimodal density of an individual sample is defined as the global density weighted by the corresponding frequency of this sample being observed in the data space. Jointly evolving and compressing fuzzy system (JECFS) (Huang et al. 2021) uses an identical mechanism as ALMMo to learn new fuzzy rules from data.

3.1.1.2 Distance criterion Dynamic evolving neural-fuzzy inference system (DENFIS) (Kasabov and Song 2002), which is another early work of evolving fuzzy systems, is the most representative model that uses distance criterion for rule generation. In DENFIS, a newly observed data sample is recognized as the focal point of a new cluster if its Euclidean distances to the existing cluster centres exceed a predefined distance threshold, suggesting that this new sample is spatially distant to the identified local models. Flexible fuzzy inference systems (FLEXFIS) (Lughofer 2008) uses a very similar approach for identifying new fuzzy rules as DENFIS. The distance threshold (the so-called *vigilance parameter*) used by FLEXFIS is normalized by the input dimensionality to avoid generating excessive clusters and fuzzy rules. Gaussian evolving fuzzy modelling system (eMG) (Lemos et al. 2011) automatically adds a new fuzzy rule to the rule base if the following two conditions are both satisfied: (1) the compatibility, which is defined as a Mahalabobis distance-based similarity measure, between the current data sample and prototypes of existing fuzzy rules falls below a predefined threshold; (2) the compatibility of the most compactable prototype to the current data sample (namely, the one with the highest compatibility value) has frequently failed to exceed the predefined compatibility threshold over a given period of time. The simplified evolving Takagi-Sugeno neuro-fuzzy (Simpl_eTS +) model (Angelov 2011) establishes a new fuzzy rule if the Euclidean distance between the current data sample and the global mean is greater or smaller than all the distances between the existing focal points and the global mean. Generalized smart evolving fuzzy systems (GS-EFS) (Lughofer et al. 2015), evolving fuzzy models (EFuMo) (Dovžan et al. 2015) and evolving possibil-istic fuzzy modelling system (ePFM) (Maciel et al. 2017) are representative fuzzy models that utilize the distance criterion, employing the Mahalanobis distance as the similarity measure. GS-EFS follows the same rule adding scheme of FLEXFIS but sets an individual distance threshold for each fuzzy rule, called *local vigilance parameter*. EFuMo normal-

izes the calculated Mahalanobis distances with the learned fuzzy covariance matrices and further requires a predefined number of consecutive input samples to satisfy the distance criterion-based rule adding condition before adding a new rule. ePFM (Maciel et al. 2017) calculates the Mahalanobis distances from the current input sample to the prototypes of existing rules. A new fuzzy rule is added to the system on the conditions that (1) the distance between the current input sample and the nearest prototype is greater than a threshold derived from a Chi-square distribution and (2) the nearest prototype has been surrounded by a sufficient amount of data samples. Parsimonious learning machine (PALM) (Ferdaus et al. 2019) measures the input and output coherence utilizing the maximal information compression index (Mitra et al. 2002) and adds a new fuzzy rule if the similarities between the hyperplanes of learned fuzzy rules and the current input sample drop below the predefined threshold. Different from evolving fuzzy systems that only use distance criterion for fuzzy rule generation, correntropy-based evolving fuzzy neural system (CEFNS) (Bao et al. 2018) and its modified version called recursive maximum correntropy-based evolving fuzzy system (RMCEFS) (Rong et al. 2019) also employ an additional error criterion. In CEFNS and RMCEFS, a new fuzzy rule will be added to the system knowledge base on condition that the Euclidean distance between the current input sample and the nearest focal point is greater than the distance threshold and the current output error does not exceed a predefined range.

3.1.1.3 Error criterion Self-organizing fuzzy neural network (SOFNN) (Leng et al. 2004, 2005) uses error criterion for fuzzy rule adding. The error criterion is derived from the geometric growing criterion (Kadirkamanathan and Niranjan 1993) and satisfies the ε-completeness of fuzzy rules (Lee 1990). Based on the error criterion, SOFNN will add a new fuzzy rule if the system output error exceeds a certain threshold given the current input sample. Both evolving fuzzy neural network (EFuNN) (Kasabov 2001) and self-constructing fuzzy neural network (SCFNN) (Lin et al. 2001) utilize the error and firing strength criteria together for adding fuzzy rules. EFuNN (Kasabov 2001) decides to add a new fuzzy rule if any of the following two conditions are satisfied: (1) the current output error is greater than the predefined error threshold, and; (2) the firing strength of the current data sample is below the firing strength threshold. However, in SCFNN (Lin et al. 2001), a more conservative fuzzy rule adding strategy is considered, which requires the two conditions used by EFuNN to be satisfied at the same time. The same strategy as SCFNN is also used in the evolving fuzzy system with self-learning/adaptive thresholds (EFS-SLAT) (Ge and Zeng 2020). Incremental fuzzy c-regression clustering-based system (InFuR) (Blazic and Skrjanc 2020) adds a new fuzzy rule to the rule base if the differences between the individual fuzzy rules' outputs with respect to the current input sample and the targeted system output all exceed the predetermined threshold. However, if the differences are all far greater than the threshold, InFuR will declare the current input sample as an outlier and discard it instead. Statistically evolving fuzzy inference system (SEFIS) (Yang et al. 2022a) uses a similar mechanism as SOFNN (Leng et al. 2004, 2005) for recruiting new fuzzy rules. Once the prediction error on the current input sample exceeds a predetermined soft threshold, a new fuzzy rule is added to the rule base of SEFIS.

3.1.1.4 Firing strength criterion The self-organizing fuzzy modified least-squares (SOFMLS) network proposed in Jesús Rubio (2009) is an early work that employs the firing strength criterion for new fuzzy rule identification. In SOFMLS, a new fuzzy rule is initialized by the current input sample if the normalized firing strengths produced by the

Table 2 Summary of fuzzy rule generation criteria used by mainstream evolving fuzzy systems

Generation criterion	Models
Density/potential	ALMMo; eClass0; eClass1; eTS; JECFS; McFIS; Simpl_eTS; SOFIS
Distance	CEFNS; DENFIS; EFuMo; eMG; ePFM; FLEXFIS; GS-EFS; PALM; RMCEFS; Simpl_eTS+
Error	EFS-SLAT; EFuNN; ESAFIS; InFuR; McFIS; SAFIS; SCFNN; SOFNN
Firing strength	ENFM; GS-EFS; GSETSK; LEOA; SEFS; SOFMLS; SPATFIS
Statistical contribution	ESAFIS; GENEFIS; PANFIS; SAFIS

fuzzy rules within the rule base are all smaller than a predefined threshold. Later works, such as generic self-evolving Takagi-Sugeno-Kang (GSETSK) fuzzy neural network (Nguyen et al. 2015), local error optimization approach for learning evolving fuzzy system (LEOA) (Ge and Zeng 2018a) and self-evolving fuzzy system (SEFS) (Ge and Zeng 2018b) also use the same firing strength criterion-based evolving scheme as SOFMLS for fuzzy rule generation. The evolving neuro-fuzzy model (ENFM) (Soleimani-B et al. 2010) adds a new rule if the maximum firing strength produced by the existing fuzzy rules on the current input sample is below the predefined threshold. In (Lughofer and Pratama 2018), an alternative rule adding mechanism combining the firing strength criterion and the degree of nonlinearity of the current model is introduced to GS-EFS. Spatio-temporal fuzzy inference system (SPATFIS) (Samanta et al. 2019) uses the maximum membership values assigned to input samples per attribute as the criterion for fuzzy rule generation. A new fuzzy rule will be added to the knowledge base if the maximum membership value assigned to one of the variables of the current input sample is smaller than the predefined threshold.

3.1.1.5 Statistical contribution criterion Sequential adaptive fuzzy inference system (SAFIS) (Rong et al. 2006) firstly proposes the statistical contribution criterion for growing the fuzzy rule base. Combining with the distance criterion, SAFIS adds a new fuzzy rule if the estimated statistical contribution of the new rule to the system outputs exceeds a predefined threshold and, at the same time, the Euclidean distance between the current input sample and the nearest focal point is greater than the distance threshold. Extended sequential adaptive fuzzy inference system (ESAFIS) (Rong et al. 2011) employs the same rule adding mechanism of SAFIS but with a simplified method to calculate the statistical contributions of fuzzy rules. In parsimonious network based on fuzzy inference system (PANFIS) (Pratama et al. 2014b) and generic evolving neuro-fuzzy inference system (GENEFIS) (Pratama et al. 2014a), the statistical contribution of each new input samples is calculated and a new fuzzy rule is initialized if the statistical contribution exceeds a predefined threshold.

A brief summary of the fuzzy rule generation criteria used by mainstream evolving fuzzy systems is given in Table 2.

3.1.2 Rule merging

Rule merging is a very useful component of the structure evolving scheme to resolve rule conflict and simplify the knowledge base. Typically, an evolving fuzzy system

decides to merge two or multiple fuzzy rules together if they are highly similar to each other. In this subsection, the commonly used rule merging mechanisms are reviewed.

EFuNN (Kasabov 2001) will aggregate multiple fuzzy rules together into one if the radius of the cluster associated with the aggregated new fuzzy rule is less than a predefined maximum radius. SOFNN (Leng et al. 2004, 2005) is equipped with a very strict rule merging mechanism that only combines these fuzzy rules with the same premise part. ENFM (Soleimani-B et al. 2010) decides to merge two fuzzy rules if the Mahalanobis distance-based similarity between the two focal points is greater than the predefined threshold. In (Lughofer et al. 2011), a new rule merging mechanism is introduced to FLEXFIS (Lughofer 2008), enabling it to merge two fuzzy rules together if the overlapping index between their membership functions is greater than the predefined threshold. eMG (Lemos et al. 2011) combines two fuzzy rules into one if the compatibility between the prototypes of the two rules exceeds the predefined threshold. PANFIS (Pratama et al. 2014b) and GENEFIS (Pratama et al. 2014a) employ the same rule merging mechanism based on the similarity degree between Gaussian membership functions of its fuzz rules. The systems will merge two fuzzy rules into one if their similarity exceeds the predefined threshold. GS-EFS (Lughofer et al. 2015; Lughofer and Pratama 2018) uses a joint criteria to examine whether two fuzzy rules need to be merged together or not. In GS-EFS, two fuzzy rules will be merged together if their overlap degree calculated based on the Mahalanobis distance between the two focal points is greater than the predefine threshold or the angle between the hyperplanes defined by their consequent parameters is greater than 90°. EFuMo (Dovžan et al. 2015) merges two fuzzy rules if any one of the following three criteria are satisfied: (1) the normalized Mahalanobis distance between two focal points is smaller than the predefined threshold; (2) the correlation coefficient calculated from the membership degrees is greater than the predefined threshold; (3) the angle between the consequent parameters is below the predefined threshold. SEFS (Ge and Zeng 2018b) employs a L^2 distance-based similarity measure calculated from membership functions as the merging criterion. If the similarity of two fuzzy rules exceeds the predefined threshold, SEFS will merge them together. The rule merging mechanism used by LEOA (Ge and Zeng 2018a) is triggered if the associated clusters of two fuzzy rules have a high overlapping level. The overlapping level is determined by the Euclidean distance between the two focal points and the areas of influences covered by the two clusters. EFS-SLAT (Ge and Zeng 2020) decides to merge two fuzzy rules if the firing strengths assigned by the two rules to each other's focal points both exceed the predefined threshold. Since the fuzzy rules of PALM (Ferdaus et al. 2019) are based on hyperplanes, two fuzzy rules will be merged if the angle and spatial proximity between their hyperplanes both fall below the predefined threshold. In such case, the fuzzy rule with the smaller support will be merged into the other one. SPATFIS (Samanta et al. 2019) keeps monitoring the similarity index between any two rules, which is calculated from the differences between their centres and spreads. Two rules will be found to be identical if the similarity index exceeds a threshold and will be merged together as one.

Another form of rule merging is to replace an existing fuzzy rule with a newly initialized new rule to avoid possible overlapping. In such cases, some of the parameters of the old rule will be inherited by the new rule. For example, the eTS model (Angelov and Filev 2004) will check the distance between the focal points of the new rule and the nearest old rule after a new fuzzy rule is initialized by the current input sample. If the distance is below a dynamical threshold calculated based on the potentials of the two focal points, the old rule will be replaced by the new one. eClass0 and eClass1 (Angelov and Zhou 2008) replace an existing fuzzy rule with the newly established one if the firing strength of the old fuzzy rule is greater than a hardcoded threshold determined by the so-called "one sigma"

Table 3 Summary of fuzzy rule pruning criteria used by mainstream evolving fuzzy systems

Pruning criterion	Models
Age	eClass0; eClass1
Utility	ALMMo; EFS-SLAT; ePFM; LEOA; Simpl_eTS +
Population	EFuMo; EFuNN; Simpl_eTS; SOFMLS
Rule importance	SOFNN; SEFIS
Rule contribution	ESAFIS; GENEFIS; McFIS; PANFIS; SAFIS; SPATFIS

condition (Duda et al. 2000). Simpl_eTS (Angelov and Filev 2005) examines the Euclidean distance between the current new focal point and the nearest previously identified focal point. The old fuzzy rule will be replaced if the distance is below the threshold. Simpl_eTS + (Angelov 2011) replaces the nearest old fuzzy rule with the newly identified fuzzy rule if its firing strength to the focal point of the new rule exceeds a predefined threshold. ALMMo (Angelov et al. 2018) uses a highly similar mechanism as Simpl_eTS (Angelov and Filev 2005) and Simpl_eTS + (Angelov 2011) by checking the *local density* of the new focal point at the cluster associated with the nearest fuzzy rule. The *local density* of a data sample is calculated locally within each individual cluster, measuring the fitness of this sample to the local model represented by the cluster. If the *local density* produced by the cluster exceeds the hardcoded threshold derived from the Chebyshev inequality, the old fuzzy rule will be replaced by the newly initialized fuzzy rule (Saw et al. 1984).

Rule merging is an effective mechanism to keep the system more compact, interpretable and adaptable. However, it is not an essential component since many existing evolving fuzzy systems are not equipped with a rule merging mechanism, and this does not stop them from achieving high-level performance on various benchmark problems and real application scenarios. For example, among the evolving fuzzy models mentioned in Sect. 3.1.1, DENFIS (Kasabov and Song 2002), CEFNS (Bao et al. 2018), McFIS (Subramanian and Suresh 2012), RMCEFS (Rong et al. 2019), SAFIS (Rong et al. 2006), ESAFIS (Rong et al. 2011), SCFNN (Lin et al. 2001), GSETSK (Nguyen et al. 2015), ePFM (Maciel et al. 2017), SOFIS (Gu and Angelov 2018a), InFuR (Blazic and Skrjanc 2020), JECFS (Huang et al. 2021) and SEFIS (Yang et al. 2022a) are not equipped with such mechanism.

3.1.3 Rule pruning

Rule pruning is a component of the structure evolving scheme to remove stale fuzzy rules from the knowledge base. The stale fuzzy rules are no longer valid and contribute little to the system outputs because they fail to describe the current data patterns. Pruning such rules effectively keeps the knowledge base of the evolving fuzzy systems clear and compact. Different evolving fuzzy systems may employ different criteria to identify stale rules. Commonly used criteria include age, utility, population, rule importance, rule contribution, etc. (Ge and Zeng 2020). A brief summary of the rule pruning criteria used by mainstream evolving fuzzy systems is presented in Table 3.

3.1.3.1 Age criterion eClass0 (Angelov and Zhou 2008) and eClass1 (Angelov and Zhou 2008) employ the age criterion for rule pruning. The age of a fuzzy rule gives the accumulated information about the time instances, at which input samples were assigned to this rule.

Each time a new input sample is assigned to a rule, the age of that rule becomes smaller, otherwise, the rule grows older. eClass0 (Angelov and Zhou 2008) and eClass1 (Angelov and Zhou 2008) will remove a particular fuzzy rule from the knowledge base if its age is one standard deviation greater than the mean of the ages of all existing rules in the system.

3.1.3.2 Utility criterion Simpl_eTS + (ANGELOV 2011), ePFM (Maciel et al. 2017), ALMMo (Angelov et al. 2018) and EFS-SLAT (Ge and Zeng 2020) utilize the utility criterion for stale rule identification. The concept of utility is defined as an average of normalized firing strengths produced by the individual fuzzy rules. In (Angelov 2011; Maciel et al. 2017; Angelov et al. 2018; Ge and Zeng 2020), a stale fuzzy rule will be identified from the knowledge base if its utility is below the pre-set tolerance. LEOA (Ge and Zeng 2018a) uses a variation of utility criterion for rule pruning, which is defined as the sum of the firing strengths from the time instance at which the fuzzy rule was initialized to the current time instance. Similarly, a fuzzy rule will be recognized as a stale one and removed from the knowledge base if the sum value is below the predefined threshold.

3.1.3.3 Population criterion The pruning mechanisms employed by EFuNN (Kasabov 2001), Simpl_eTS (Angelov and Filev 2005), EFuMo (Dovžan et al. 2015) and SOFMLS (Jesús Rubio 2009) are based on the population criterion. EFuNN (Kasabov 2001) removes a fuzzy rule from the knowledge base if this rule fails to receive enough data samples after a user-defined period of time after it was initialized. During operation, Simpl_eTS (Angelov and Filev 2005) consistently monitors the support (namely, population) of each fuzzy rule. A stale rule will be identified and removed from the knowledge base if its support is less than 1% of the overall number of observed samples. EFuMo (Dovžan et al. 2015) prunes a stale rule if the average number of data samples assigned to this rule per instance is lower than a predefined percentage of the mean value of the entire rule base or this rule fails to get a sufficient amount of data samples after a certain period of time. SOFMLS (Jesús Rubio 2009) repeatedly examines the support of every fuzzy rule in the knowledge base for every a predetermined number of instances and will remove a rule if its support is smaller than the predefined threshold.

3.1.3.4 Rule importance criterion SOFNN (Leng et al. 2004, 2005) prunes fuzzy rules based on their importance. It will evaluate the importance of every fuzzy rule to system performance by removing it temporarily and calculating the changes of root mean squared error (RMSE) of its predictions. Based on this, SOFNN identifies the rules that only incurs minor changes to its prediction accuracy. Such fuzzy rules will be pruned from the knowledge base if the prediction error of SOFNN is below the predefined value without them. SEFIS (Yang et al. 2022a) monitors its prediction error on every input sample consistently and prunes the fuzzy rule with the nearest prototype to the current input if the prediction error drops below the predetermined soft threshold, under the assumption that other fuzzy rules are sufficient for approximation.

3.1.3.5 Rule contribution criterion The rule pruning mechanisms of SAFIS (Rong et al. 2006), ESAFIS (Rong et al. 2011), McFIS (Subramanian and Suresh 2012), PANFIS (Pratama et al. 2014b), GENEFIS (Pratama et al. 2014a) and SPATFIS (Samanta et al. 2019) are based on the rule contribution criterion. These evolving fuzzy systems use different approaches to evaluate the statistical contributions of the fuzzy rules to system output, but they follow the same procedure for rule pruning. During the learning process, the statis-

tical contribution of every fuzzy rule will be consistently monitored, and a certain rule will be pruned if its contribution falls below a predefined threshold.

A well-designed rule pruning mechanism can effectively enhance the capability of an evolving fuzzy system to self-adapt to the unfamiliar patterns of data streams. It plays a key role in rule base simplification, effectively improving the computation- and memory-efficiency of the system. However, same as rule merging, rule pruning is also not an essential component for an evolving fuzzy system to perform well in real application scenarios. For example, the following mainstream evolving fuzzy systems are not equipped with a rule pruning mechanism: eTS (Angelov and Filev 2004), DENFIS (Kasabov and Song 2002), eMG (Lemos et al. 2011), FLEXFIS (Lughofer 2008), SOFIS (Gu and Angelov 2018a), GS-EFS (Lughofer et al. 2015; Lughofer and Pratama 2018), CEFNS (Bao et al. 2018), RMCEFS (Rong et al. 2019), SCFNN (Lin et al. 2001), GSETSK (Nguyen et al. 2015), PALM (Ferdaus et al. 2019), SEFS (Ge and Zeng 2018b), ENFM (Soleimani-B et al. 2010), InFuR (Blazic and Skrjanc 2020) and JECFS (Huang et al. 2021).

3.1.4 Rule splitting

Rule splitting is another component of the structure evolving scheme to assist evolving fuzzy systems to build finer system structure and achieve finer partition of the data space. However, compared with rule merging and rule pruning, rule splitting is a relatively new concept and is rarely used by the existing evolving fuzzy models.

Among the mainstream evolving fuzzy systems mentioned in Table 1, EFuMo (Dovžan et al. 2015) and GS-EFS (Lughofer et al. 2015; Lughofer and Pratama 2018) are the only two equipped with such mechanism. EFuMo (Dovžan et al. 2015) determines whether to split a fuzzy rule or not based on its support and mean relative error, which is calculated based on the weighted average prediction error of this local model over time. EFuMo (Dovžan et al. 2015) splits a fuzzy rule into two if its support is greater than a predefined number and its mean relative error is larger than the threshold. In (Lughofer et al. 2018), a similar but more advanced rule splitting mechanism is introduced to GS-EFS (Lughofer et al. 2015; Lughofer and Pratama 2018). Without using externally controlled thresholds as EFuMo (Dovžan et al. 2015) does, GS-EFS (Lughofer et al. 2018) decides to split the latest updated fuzzy rule if its support and weighted average prediction error both are one or two standard deviation(s) greater than the mean values of all existing fuzzy rules.

3.1.5 Input attribute reduction

Apart from the aforementioned mainstream evolving schemes (namely, generation, merging, pruning and splitting), there are also a few evolving fuzzy systems equipped with input attribute reduction schemes to remove redundant attributes, thereby improving the prediction performance and increasing the computational efficiency. The best-known examples that are equipped with input attribute reduction schemes include Simpl_eTS + (Angelov 2011), GENEFIS (Pratama et al. 2014a), ALMMo (Angelov et al. 2018) and JECFS (Huang et al. 2021). The first three evolving fuzzy systems prune the input attributes based on their importance to the system outputs. However, Simpl_eTS + (Angelov 2011) and ALMMo (Angelov et al. 2018) perform online input selection for each individual fuzzy rule separately, such that an input attribute previously removed for a particular rule will still be considered by other rules. In contrast, the online input selection of GENEFIS

(Pratama et al. 2014a) is performed on the system-level. Once an input attribute is pruned, the dimensionality of the system inputs is reduced by 1. JECFS (Huang et al. 2021), on the other hand, attempts to learn a more compact, simpler fuzzy rule base from data streams by using very sparse random projection matrices (Li et al. 2006) to compress the dimensionality of the input data space and ease the effect of "the curse of dimensionality".

3.2 Parameter learning schemes

The premise part in the form of focal points of the fuzzy rules within evolving fuzzy systems is typically learned during the structure evolving process through online clustering the data streams. As the structure evolving schemes used by different evolving fuzzy systems vary in ways that fuzzy rules are identified, merged, pruned as well as split, the premise parts learned from streaming data can also be different, leading to the highly diverse behaviours of evolving fuzzy systems. However, the vast majority of first-order evolving fuzzy systems, which are built upon first-order fuzzy rules, in the literature utilize recursive least square (RLS)-based techniques for consequent parameter learning. The fuzzily weighted RLS (FWRLS) algorithm (Angelov and Filev 2004) is the most popular one used by a wide variety of first-order evolving models as "golden standard" (Lughofer 2011b, 2021; Angelov and Gu 2018b; Škrjanc et al. 2019). Other consequent parameter updating schemes used by first-order evolving fuzzy systems include weighted RLS (Kasabov and Song 2002), extended Kalman filter (Rong et al. 2006), stable gradient descent (Jesús Rubio and Bouchachia 2017), adaptive maximum correntropy extended Kalman filter (Yang et al. 2022a), etc. In this section, the algorithmic procedure of FWRLS is presented for illustrating the key concept.

The consequent parameters can be learned with the FWRLS algorithm either globally or locally with the objective of minimizing the mean squared error of system outputs:

$$E_k = \frac{1}{k} \sum_{i=1}^{k} \|y_i - \hat{y}_i\|^2 \tag{14}$$

where $\hat{y}_i$ is the system output at the ith time instance; y_i is the corresponding targeted value.

With the global learning approach, the consequent parameters of all first-order fuzzy rules within the knowledge base are updated together simultaneously using Eqs. (15) and (16) (Angelov and Filev 2004; Angelov and Gu 2018b):

$$\Theta_k \leftarrow \Theta_{k-1} - \frac{\Theta_{k-1} \bar{\mathbf{x}}_k^T \bar{\mathbf{x}}_k \Theta_{k-1}}{1 + \bar{\mathbf{x}}_k^T \Theta_{k-1} \bar{\mathbf{x}}_k} \tag{15}$$

$$\mathbf{a}_k \leftarrow \mathbf{a}_{k-1} + \Theta_k \bar{\mathbf{x}}_k \left(y_k - \bar{\mathbf{x}}_k^T \mathbf{a}_{k-1} \right) \tag{16}$$

where $\mathbf{x}_k$ is the input sample at the kth time instance; y_k is the corresponding targeted system output; $\bar{\mathbf{x}}_k$ is a $(N+1)L \times 1$ dimensional vector, $\bar{\mathbf{x}}_k = \left[\bar{\lambda}_{k,1} \bar{\mathbf{x}}_k^T, \bar{\lambda}_{k,2} \bar{\mathbf{x}}_k^T, \dots, \bar{\lambda}_{k,L} \bar{\mathbf{x}}_k^T \right]^T$; $\bar{\mathbf{x}}_k = \left[1, \mathbf{x}_k^T \right]^T$; $\mathbf{a}_k = \left[\mathbf{a}_{k,1}^T, \mathbf{a}_{k,2}^T, \dots, \mathbf{a}_{k,L}^T \right]^T$; $\mathbf{a}_{k,i}$ is the consequent parameter vector of the ith rule $\mathbf{R}_i$ at the kth time instance; Θ_k is the $(N+1)L \times (N+1)L$ dimensional covariance matrix calculated globally.

With the local learning approach, the consequent parameters of each individual fuzzy rule are updated in parallel, independently using Eqs. (17) and (18) (Angelov and Filev 2004; Angelov and Gu 2018b):

$$\boldsymbol{\Theta}_{k,i} \leftarrow \boldsymbol{\Theta}_{k-1,i} - \frac{\overline{\lambda}_{k,i}\boldsymbol{\Theta}_{k-1,i}\overline{\boldsymbol{x}}_k^T\overline{\boldsymbol{x}}_k\boldsymbol{\Theta}_{k-1,i}}{1 + \overline{\lambda}_{k,i}\overline{\boldsymbol{x}}_k^T\boldsymbol{\Theta}_{k-1,i}\overline{\boldsymbol{x}}_k} \tag{17}$$

$$\boldsymbol{a}_{k,i} \leftarrow \boldsymbol{a}_{k-1,i} + \overline{\lambda}_{k,i}\boldsymbol{\Theta}_{k,i}\overline{\boldsymbol{x}}_k\left(y_k - \overline{\boldsymbol{x}}_k^T\boldsymbol{a}_{k-1,i}\right) \tag{18}$$

where $\boldsymbol{\Theta}_{k,i}$ is the $(N + 1) \times (N + 1)$ dimensional covariance matrix associated with $\mathbf{R}_i$.

In general, the local learning approach with the FWRLS algorithm is more computationally efficient than the global learning approach thanks to the recursive updating of covariance matrix per rule. In addition, the learned consequent parameters by the local learning approach are less influenced by the structure evolution of the fuzzy model.

3.3 Performance comparison and analysis

In this subsection, numerical results on a number of popular benchmark datasets are presented for performance comparison between different evolving fuzzy systems. However, as the performances of evolving fuzzy systems usually are sensitive to parameter settings, the results are obtained directly from the literature (Pratama et al. 2014b; Angelov et al. 2018; Ge and Zeng 2018a, b, 2020; Rong et al. 2018, 2019; Samanta et al. 2019; Gu et al. 2021a) for a fair comparison.

Mackey–Glass chaotic time series prediction problem (Mackey and Glass 1977) is used as the first numerical example, which is one of the most widely used benchmark datasets (Kasabov and Song 2002; Rong et al. 2006, 2011). This time series is created using the Mackey–Glass time-delay differential equation. Performances of different evolving fuzzy systems reported in the literature (Ge and Zeng 2018a, b; Rong et al. 2019; Samanta et al. 2019) are listed in Table 4 in terms of non-dimensional error index (*NDEI*), number of fuzzy rules (#(*Rule*)) and training time consumptions (t_{exe}).

Secondly, Delta Ailerons dataset from the KEEL-dataset repository (Alcalá-Fdez et al. 2011a)[1] is employed as the second example, which is obtained from the tasks of controlling the ailerons of a F16 aircraft. Performances of different evolving fuzzy systems reported in the literature (Ge and Zeng 2020; Rong et al. 2018, 2019; Gu et al. 2021a) on Delta Ailerons dataset are listed in Table 5 in terms of root mean square error (*RMSE*), #(*Rule*) and t_{exe}.

Then, the S&P 500 close price prediction problem is used for experimental comparison (Pratama et al. 2014b). The daily close prices of S&P500 are collected from the Yahoo! Finance website,[2] ranging from 03.01.1950 to 12.03.2009 (60 years). Performance comparison between different evolving fuzzy systems on this problem is tabulated in Table 6 in terms of *NDEI* and #(*Rule*). The reported results are obtained directly from (Pratama et al. 2014b; Angelov et al. 2018; Ge and Zeng 2020).

It can be observed from Tables 4, 5 and 6 that the performances of evolving fuzzy systems vary a lot on different problems. For example, eTS and ALMMo produced the

---

[1] Available at: https://sci2s.ugr.es/keel/datasets.php.

[2] Available at: https://uk.finance.yahoo.com/.

Table 4 Performance comparison between different evolving fuzzy systems on Mackey-Glass time series

Model	NDEI	#(Rule)	t_{exe}
ALMMo	0.402	9	0.5
CEFNS	0.264	5	0.4
DENFIS	0.278	58	–
EFuMo	0.139	41	–
EFS-SLAT	0.114	8	–
ESAFIS	0.296	6	5.8
eTS	0.381	9	3.1
GENEFIS	0.280	19	4.5
GSETSK	0.347	19	–
LEOA	0.248	42	144.8
McFIS	0.250	10	8.2
PANFIS	0.301	19	4.5
RMCEFS	0.117	5	0.3
SAFIS	0.376	6	–
SEFS	0.129	4	0.4
Simpl_eTS	0.394	11	4.4
Simpl_eTS +	0.375	15	–
SPATFIS	0.279	30	–

Table 5 Performance comparison between different evolving fuzzy systems on Delta Ailerons

Model	RMSE	#(Rule)	t_{exe}
ALMMo	0.0513	10	0.4
CEFNS	0.0502	3	0.3
DENFIS	0.0497	11	–
EFS-SLAT	0.0412	8	–
ESAFIS	0.0506	13	12.4
eTS	0.0513	4	2.1
McFIS	0.0509	15	–
RMCEFS	0.0498	2	0.2
SAFIS	0.0549	14	6.8
SEFS	0.0476	7	0.4
Simpl_eTS	0.0512	4	2.0

most accurate predictions on predicting the S&P 500 close price, but were outperformed by EFS-SLAT on Mackey-Glass time series. In fact, the prediction performance of a particular evolving fuzzy system is determined by many different factors in relation to the system specification, such as fuzzy rule types, structural evolving schemes, parameter learning schemes, membership function types, etc. An EIS may also behave very differently for different problems depending on their natures and the experimental settings.

Table 6 Performance comparison between different evolving fuzzy systems on S&P 500 close price

Model	NDEI	#(Rule)
ALMMo	0.013	8
DENFIS	0.020	6
EFuNN	0.154	114.3
EFS-SLAT	0.016	23
eTS	0.015	14
GENEFIS	0.070	2
LEOA	0.123	52
PANFIS	0.090	4
SEFS	0.018	2
Simpl_eTS	0.045	7
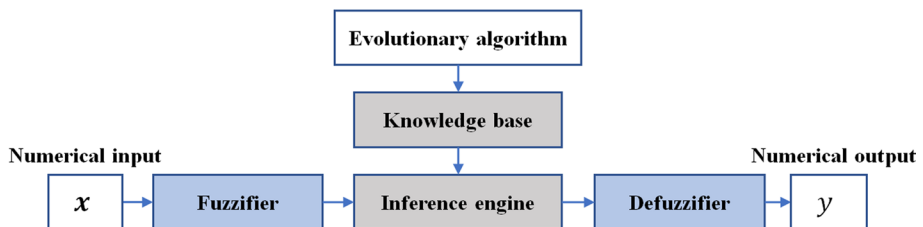
Fig. 3 Standard structure of evolutionary fuzzy systems (Herrera 2008; Elhag et al. 2019)

4 Evolutionary fuzzy systems

As mentioned in Sect. 2, Zadeh-Mamdani type fuzzy systems (Zadeh 1973; Mamdani and Assilian 1975) were introduced in 1970s, and Takagi-Sugeno-Kang type fuzzy systems (Takagi and Sugeno 1985; Sugeno and Kang 1988) were introduced in 1980s. Theoretical basis of evolutionary computation was established around the same period of time (Holland 1975), while the pioneering works of evolutionary fuzzy systems, as a hybridization between fuzzy systems and evolutionary computation, firstly appeared in early 1990s (Karr 1991; Valenzuela-Rendon 1991; Thrift 1991; Pham and Karaboga 1991).

Nature-inspired evolutionary algorithms, e.g., genetic algorithms, genetic programming, particle swarm optimization, ant colony optimization, etc., are generally applied to global optimization. Evolutionary fuzzy systems is a class of fuzzy system, which can be either Zadeh-Mamdani type or Takagi-Sugeno-Kang type, with the knowledge bases learned or tuned by evolutionary algorithms, as depicted in Fig. 3 (Herrera 2008; Fazzolari et al. 2013; Fernandez et al. 2019; Elhag et al. 2019). The design of a fuzzy system can be considered as a search task from available observations for the suitable solutions that can best approximate the problem based on the given performance metric (Fernandez et al. 2019). In evolutionary fuzzy systems, such search task is performed by population-based evolutionary and bio-inspired algorithms (thanks to their strong ability in searching for near-optimal solutions in a wide range of problem spaces). Hence, evolutionary fuzzy systems have demonstrated excellent performance in different application scenarios for handling classification and regression problems, and is currently a popular research area (Fazzolari et al. 2013; Elhag et al. 2019).
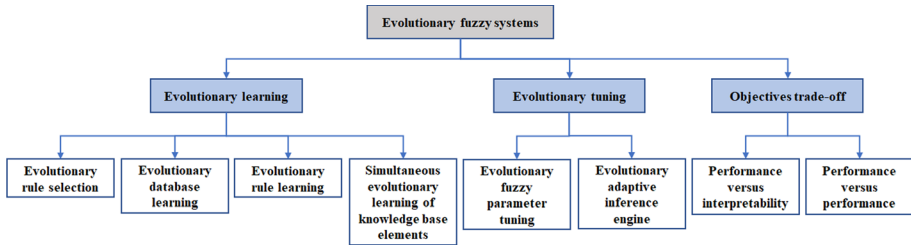
Fig. 4 Taxonomy of mainstream evolutionary fuzzy systems (Fazzolari et al. 2013; Fernández et al. 2015; Fernandez et al. 2019)

Evolutionary fuzzy systems employ evolutionary algorithms to optimize the model structure and parameters in terms of objective functions based on accuracy, interpretability or a combination of both (Fernandez et al. 2019). In general, evolutionary algorithms can be employed to learn the knowledge base of the fuzzy system (namely, evolutionary learning) or to tune a given knowledge base (namely, evolutionary tuning) (Herrera 2008). Multi-objective evolutionary algorithms can also be utilized to balance the prediction precision and interpretability of the fuzzy models (Fazzolari et al. 2013). According to the different elements of fuzzy models developed by evolutionary algorithms, a taxonomy of mainstream evolutionary fuzzy systems is given in Fig. 4. In the rest of this section, the representative approaches for developing evolutionary fuzzy systems are summarized in accordance with Fig. 4. For a more detailed literature review, interested readers are referred to the survey papers (Herrera 2008; Cordón 2011; Fazzolari et al. 2013; Fernández et al. 2015; Fernandez et al. 2019). Additional materials can be found on the website.[3]

4.1 Evolutionary learning

The aim of evolutionary learning is to learn the knowledge bases of the fuzzy systems from data with the help of evolutionary algorithms. There are four approaches for evolutionary learning.

4.1.1 Evolutionary rule selection

The main purpose of this approach is to remove useless, redundant, erroneous and/or conflictive fuzzy rules from the original knowledge base of a candidate fuzzy system, resulting in a more compact, optimized knowledge base and greater prediction precision (Ishibuchi et al. 1995, 1997; Cordón and Herrera 2000). This is very similar to the rule pruning mechanism of evolving fuzzy systems, but is implemented by evolutionary algorithms.

4.1.2 Evolutionary data base learning

This approach aims to learn the database of knowledge base, which include membership functions, fuzzy partition granularity, scaling functions. This can be achieved through two different methods. One is to use a measure to directly evaluate the quality of the generated fuzzy parameters. The other is to evaluate the quality of the entire knowledge base derived

---

[3] Available at: http://sci2s.ugr.es/moefs-review/

from the generated fuzzy parameters instead. Examples of evolutionary database learning can be found in (Park et al. 1994; Cordón et al. 2001b, c), where the second approach was implemented.

4.1.3 Evolutionary rule learning

This approach is designed to learn fuzzy rules by utilizing evolutionary algorithms with a predefined database. The earliest example of this approach was presented in Thrift (1991). In (Ishibuchi et al. 1999), individual fuzzy rules are generated by genetic algorithms for classification tasks, where the linguistic terms and membership functions are fixed a priori. The work presented in Rodríguez-Fdez et al. 2016a utilizes an ad hoc method to estimate the fuzzy partition granularity from data and then uses the evolutionary algorithm to learn the fuzzy rules. This work is further extended in Rodríguez-Fdez et al. 2016b for big data.

4.1.4 Simultaneous evolutionary learning of knowledge base elements

This approach learns both the fuzzy rules and data base together from data using evolutionary algorithms (Homaifar and McCormick 1995; Shi et al. 1999). However, this means that the searching space will be much larger, and the convergence speed will be lower. A hybrid approach combining the evolutionary learning of adaptive inference engine and knowledge base elements is proposed in Márquez et al. (2007).

4.2 Evolutionary tuning

Evolutionary tuning aims to improve the performances of existing fuzzy systems by adjusting the parameters of the knowledge bases or inference engines using evolutionary algorithms.

4.2.1 Evolutionary knowledge base tuning

This approach optimizes the knowledge base of a learned fuzzy system by adjusting its membership function parameters through an evolutionary tuning process (Angelov and Guthke 1997; Casillas et al. 2005). Some works (Angelov 1999, 2000; Alcalá et al. 2007a; Alcalá-Fdez et al. 2011b) combine knowledge base tuning with rule selection to achieve higher performance. In the works presented in (Gacto et al. 2009; Alcalá et al. 2007b; Pulkkinen and Koivisto 2008), multi-objective evolutionary algorithms are employed to perform membership function parameter tuning and rule selection simultaneously in order to achieve a trade-off between accuracy and complexity.

4.2.2 Evolutionary adaptive inference engine

The main aim of this approach is to improve the system performance by adjusting the fuzzy inferencing scheme. An evolutionary adaptive inference engine can be implemented in two different ways. One way is to introduce adaptive parameters, which are tuned by evolutionary algorithms, to the inference system such that higher cooperation among the fuzzy rules can be achieved (Alcalá-Fdez et al. 2007). The other way is to introduce adjustable weights to the defuzzifier and use evolutionary algorithms to tune these weights (Kim et al. 2002).

Table 7 Performance comparison between ALMMo and PSO-ALMMo

Dataset	Model	$NDEI$	#($Rule$)	t_{exe}
Mackey-Glass	ALMMo	0.4320	8	0.5
	PSO-ALMMo	0.1910		314.3
Delta ailerons	ALMMo	0.5499	10	0.5
	PSO-ALMMo	0.5351		362.4
S&P 500	ALMMo	0.0147	6	3.6
	PSO-ALMMo	0.0146		2783.3

4.3 Objective trade-off

Although the primary goal of introducing evolutionary algorithms to the design of a fuzzy system is to improve its prediction precision, it is also possible to consider different criteria such as interpretability, stability, robustness, computational efficiency during the system design process by utilizing multi-objective evolutionary algorithms.

4.3.1 Performance versus interpretability

As aforementioned, prediction precision/accuracy is the key criterion considered during the system design. Indeed, the similarity between the responses of the real system and the fuzzy model should be as high as possible in real application scenarios (Gacto et al. 2011). The interpretability of the fuzzy model is also very important considering the current trend towards explainable AI. However, interpretability is a more subjective criterion and there is no standard approach to quantify it. There are several system complexity-based factors considered to be relevant to the model interpretability (Gacto et al. 2011; Rudzi 2016), for example, the number of input variables, the number of fuzzy rules, the number of membership functions, etc. In addition, there are semantic-based factors to be considered, such as incoherence, distinguishability and rule relevance (Rey et al. 2017). By balancing between prediction accuracy and model interpretability, a multi-objective evolutionary fuzzy system can achieve adequately high prediction performance with a clear and compact knowledge base. Examples of such approaches can be found in (Alcalá et al. 2007b; Gacto et al. 2009; Fazzolari et al. 2014; Rudzi 2016; Rey et al. 2017; Gorzałczany and Rudziński 2017).

4.3.2 Performance versus performance

During the design of industrial control systems, there are often a wider variety of criteria apart from prediction precision and model interpretability needed to be considered, which may include robustness, time efficiency, stability (Elhag et al. 2019). To achieve the trade-off between different criteria, multi-objective evolutionary algorithms are usually employed in the system design. One can use multi-objective evolutionary algorithms to learn the knowledge base and/or its components from data to construct the fuzzy system, but a more commonly used approach is to tune the structure and parameters of an existing fuzzy model (Fernández et al. 2015; Fernandez et al. 2019).

4.4 Performance demonstration and analysis

In this subsection, numerical results obtained by ALMMo and particle swarm optimized ALMMo (PSO-ALMMo) on the three benchmark datasets used before are presented as an example. PSO-ALMMo is the evolutionary version of ALMMo. It utilizes a particle swarm optimization algorithm (Eberhart and Kennedy 1995) to simultaneously optimize the premise and consequent parameters learned by ALMMo from streaming data based on historical observations, namely, evolutionary tuning. However, in order to achieve this, PSO-ALMMo requires historical data to be stored in system memory, and its optimization process has to be conducted offline. Numerical results by ALMMo and PSO-ALMMo obtained from Gu et al. 2021a are tabulated by Table 7 as performance comparison. Note that ALMMo in Gu et al. 2021a utilizes Gaussian type membership function instead of the Cauchy membership function as used in the original version (Angelov et al. 2018), causing differences in terms of prediction precision, system complexity and computational efficiency.

Table 7 shows that PSO-ALMMo is able to produce more accurate predictions than ALMMo thanks to the iterative parameter optimization process. However, as the parameter optimization process can take a great amount of time before converging to a locally optimal solution depending on the complexity and dimensionality of the searching space, this process has to be performed offline. As a result, PSO-ALMMo is limited to offline application scenarios despite of its greater precision.

5 Reinforcement learning-based fuzzy systems

Reinforcement learning is a machine learning paradigm for solving decision-making problems in Markovian processes (Sutton and Barto 2018). Reinforcement learning problems are essentially closed-loop problems in the sense that the actions of the learning system (also called agent) influence its later inputs. In a typical reinforcement learning framework, an agent learns to achieve a goal by interacting with the environment, which is defined in the form of a Markov decision process. The agent gets either rewards or penalties for the actions it performs, and its main goal is to maximize the long-term reward (Huang et al. 2020).

Reinforcement learning differs from supervised learning as it does not require input–output training data to be explicitly presented during the training process. Instead, the learning system is expected to figure out the solution yielding the maximum return in the form of rewards as the result of its actions. To begin with, the agent only needs to be equipped with prior knowledge of possible actions and reward policy. Thanks to the capability of discovering optimal solutions through interactions with the environment, reinforcement learning has been extensively researched in recent years and widely used in the domains such as gameplay (Silver et al. 2016), robotics (Gu et al. 2017).

Reinforcement learning is closely associated with adaptive control and optimal control (Lewis and Vrabie 2012). Pioneering works of utilizing reinforcement learning in fuzzy control system design appeared in early 1990s, which include approximate reasoning-based intelligent control (ARIC) system (Berenji 1992), generalized approximate reasoning-based intelligent control (GARIC) system (Berenji and Khedkar 1992), reinforcement neural-network-based fuzzy logic control system (RNN-FLCS) (Lin and George Lee 1994) and reinforcement fuzzy adaptive learning control network (RFALCON) (Lin and

Lin 1996). These control systems utilize two separate sub-models to perform Actor–Critic learning (Konda and Tsitsiklis 2000) and, thus, their system structures are complex. The system structures of ARIC and GARIC are both fixed, and they use reinforcement learning techniques to tune the parameters only. In contrast, RNN-FLCS and RFALCON are capable of self-developing the system structure and parameters, but they still require experts to determine the number and type of membership function for each input and output variable. In (Jouffe 1998), two fuzzy reinforcement learning methods, fuzzy actor-critic learning (FACL) and fuzzy Q-learning (FQL) are proposed for tuning the consequent parameters of fuzzy controller. A fuzzy Actor–Critic reinforcement learning network (FACRLN) is proposed in Wang et al. (2007). FACRLN uses a single fuzzy radial basis function neural network to approximate both the action value function of the Actor and the state value function of the Critic simultaneously, thus, the system complexity is largely reduced. In (Juang and Lu 2009), an ant colony optimization-based algorithm combined with FQL is proposed for determining the consequent parameters of fuzzy inference systems. By treating the combined return value of a series of actions as the fitness value to be maximized, a particle swarm reinforcement learning method is presented in Hein et al. (2017) to learn the best policy represented by fuzzy rules. Since the majority of existing fuzzy reinforcement learning methods are implemented on the basis of (fuzzy) neural networks with very limited interpretability, an interpretable reinforcement learning scheme is proposed in Huang et al. (2020), where the learned policy can be expressed as human-intelligible IF-THEN rules and the value function is approximated through the AnYa type fuzzy rule-based system.

6 Comparison between evolving, evolutionary and reinforcement learning-based fuzzy systems

Evolutionary fuzzy systems and evolving fuzzy systems are closely related. In some works, EFSs is short for evolutionary fuzzy system (Fazzolari et al. 2013; Fernández et al. 2015; Fernandez et al. 2019), and in many other works, EFSs is the abbreviation of evolving fuzzy systems (Škrjanc et al. 2019; Campos Souza 2020; Leite et al. 2020). Few early works of evolutionary fuzzy systems characterize themselves as evolving (Angelov 1999, 2000). Despite of the similarity, there are some key differences between the two classes of fuzzy systems.

Evolutionary algorithms play an instrumental role in the design of evolutionary fuzzy systems. The evolutionary learning process is performed by the operators that mimic the natural evolutionary phenomena such as chromosomes crossover, mutation, selection and reproduction, parents and off-springs. Since evolutionary algorithms search for the optimal solutions in the problem spaces through an iterative process, evolutionary fuzzy systems usually can achieve strong, robust performance in problems with complex nature. Nevertheless, the training processes of evolutionary fuzzy systems are limited to offline due to the requirement of iterative computation. In order to obtain the (nearly) optimal solution, evolutionary fuzzy systems require all the training data needs 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.

In contrast, evolving fuzzy systems are designed to gradually self-adapt the system structure and parameters to follow the concept shifts and drifts of underly patterns of the data streams. The vast majority of evolving fuzzy systems learn from streaming data in a single-pass, non-iterative manner, and they stress the ability to react rapidly to the

dynamically changing data patterns in nonstationary environments. This is certainly correct considering the targeted application scenarios and is of great importance to the success of evolving fuzzy systems (Gu et al. 2021a). However, such learning behaviours can sometimes lead to poor global prediction accuracy and the so-called "unlearning effect" (Ge and Zeng 2018a) as the models tend to fit the more recently arrived data well but fit historical data poorly.

On the other hand, reinforcement learning is generally less effective than supervised learning if there exists sufficient training data. It assumes that the world is Markovian and requires prior knowledge about the environment and possible actions of the agent. Huge amounts of time and computational resources are usually needed before the reinforcement learning system finding out the optimal solution. In addition, it sometimes can be very challenging to define the reward policies for some real-world problems, such as autonomous driving (Kiran et al. 2021). Thus, to date, reinforcement learning-based fuzzy systems are mostly implemented for control problems in the robotics and industrial automation domains where prior knowledge is sufficient to quantify the states of environments and determine the possible actions of agents (Yung and Ye 1999; Lin and Jou 2000; Fathinezhad et al. 2016).

Table 8 presents a brief comparison between the three classes of fuzzy systems discussed from the following six different aspects: (1) learning mode (online/offline); (2) requirement for supervision (input–output pairs); (3) structure and parameter learning (self-developing or pre-fixed by user); (4) optimization of the structure and/or parameters during training (yes/no); (5) computational complexity (high/low), and; (6) need for prior knowledge (weak to strong).

7 Applications of fuzzy systems

Examples of real-world applications of the three classes of fuzzy systems are listed in Table 9, which cover a wide range of areas including agriculture, communication, computing, healthcare, finance, remote sensing, etc. This demonstrates the flexibility and wide applicability of evolving, evolutionary and reinforcement learning-based fuzzy systems, as powerful tools for handling real-world problems with different nature. Due to the very large number of research publications containing applications of fuzzy systems, which is also an empirical evidence about the importance of these methods, it is impossible to cover all of them in this overview paper. More detailed reviews of recent applications of these methods in real-world problems can be found from (Škrjanc et al. 2019; Campos Souza 2020; Leite et al. 2020; Fernández et al. 2015; Fernandez et al. 2019).

8 Stock price prediction-a case study

To demonstrate the utility of fuzzy rule-based systems in real-world application scenarios, a case study of predicting the stock price of Walmart Inc. (WMT) is presented. The daily high, low, open and close prices of WMT are acquired from the Yahoo! Finance website,[4] ranging from 01.01.2000 to 01.01.2021 (20 years), 5284 samples in total. In

---

[4] Available at: https://uk.finance.yahoo.com/

Table 8 Comparison between evolving, evolutionary and reinforcement learning-based fuzzy systems

Fuzzy systems	Learning mode	Requirement for supervision	Structure and parameters	Optimization during training	Computational complexity	Need for prior knowledge
Evolving	Online	Yes	Self-developing	No	Low	Weak
Evolutionary	Offline	Yes	Partially predefined	Yes	High	Medium
Reinforcement learning-based	Offline	No	Partially predefined	Yes	High	Strong

Table 9 Real-world applications of evolving, evolutionary and reinforcement learning-based fuzzy systems

Application	Fuzzy system		
	Evolving	Evolutionary	Reinforcement learning-based
Agriculture	Alghamdi et al. (2019), Qaddoum et al. (2013)	Mohamed and Hameed (2018)	
Autonomous driving/driving assistance	Soares et al. (2019a), Andonovski et al. (2021a), Škrjanc et al. (2018a), Zhang et al. (2019a), Soares et al. (2019b)	Onieva et al. (2015)	Dai et al. (2005), Wu, et al. (2021)
Biomedical diagnosis	Lekkas and Mikhailov (2010), Stirling et al. (2020), Soares et al. (2020), P. V. de C. Souza and E. Lughofer (2020), Novitasari et al. (2019), Liu et al. (2016), Jesús Rubio (2014)	Pe and Sipper (1999), Carmona et al. (2011)	Kukker and Sharma (2020)
Energy system	Severiano et al. (2021), Chang et al. (2011), Li et al. (2014), Andonovski et al. (2021b), Andonovski et al. (2016)	Safari et al. (2013), Hinojosa and Hoese (2010), Musilek et al. (2015), Blažič et al. (2021)	Alfaverh et al. (2020), Yu and Dexter (2010)
Fault diagnosis and detection	Pratama et al. (2020a), Silva et al. (2020), Costa et al. (2015), Lemos et al. (2013), Du et al. (2020), Camargos et al. (2020)	Serdio et al. (2014), Razavi-Far et al. (2009)	Malik et al. (2020)
Financial time series prediction	Rajab and Sharma (2019), Esfahanipour and Aghamiri (2010), Gu et al. (2016), Rodrigues Júnior and Oliveira Serra (2020), Yong et al. (2018), Yu and Zhang (2005), Garcia et al. (2019)	Mehmanpazir and Asadi (2017), Feng and Chou (2012), Hadavandi et al. (2010), Cao et al. (2020), Sanz et al. (2015)	
Human action recognition	Sargano et al. (2020), Andreu and Angelov (2010), Iglesias et al. (2010)	Fahim et al. (2013)	
Industrial control	Precup et al. 2020; Leite et al. 2015; Mendes et al. 2017; Yang et al. 2021a; Škrjanc et al. 2014; Precup et al. 2014; Škrjanc 2021	Choi et al. (2015), Tsourveloudis (2010)	Zhang et al. (2019b)
Intrusion detection (computer networks)	Decker et al. 2020; Iglesias et al. 2013a; Iglesias et al. 2012; Škrjanc et al. 2018b	Elhag et al. (2019), Abadeh et al. (2011), Elhag et al. (2015), Tsang et al. (2007), Naik et al. (2018)	

Table 9 (continued)

Application	Fuzzy system		Reinforcement learning-based
	Evolving	Evolutionary	
Remote sensing	Gu et al. 2022; Gu et al. 2018	Stavrakoudis et al. 2011, Mylonas et al. 2013	
Robotics	Sadeghi-Tehran and Angelov 2011; Gu et al. 2020a	Chatterjee et al. 2005, Juang et al. 2020	Zhou and Meng 2003, Ye et al. 2003, Goharimanesh et al. 2020
Wastewater treatment	Dovžan et al. 2015; Dovžan and Škrjanc 2011	Han et al. 2018, Qiao et al. 2018	
Weather forecasting	Soares et al. 2018a, Soares et al. 2018b, Saputro et al. 2016, Leite and Škrjanc 2019	Azad et al. 2020	Malik and Yadav 2021
Wireless network	Huang et al. 2004	Prauzek et al. 2016	Razavi et al. 2010, Muñoz et al. 2015, Khalid et al. 2021, Yang et al. 2021b

(a) Prediction result

(b) Stepwise prediction error

(c) Average squared prediction error

(d) Rule evolution
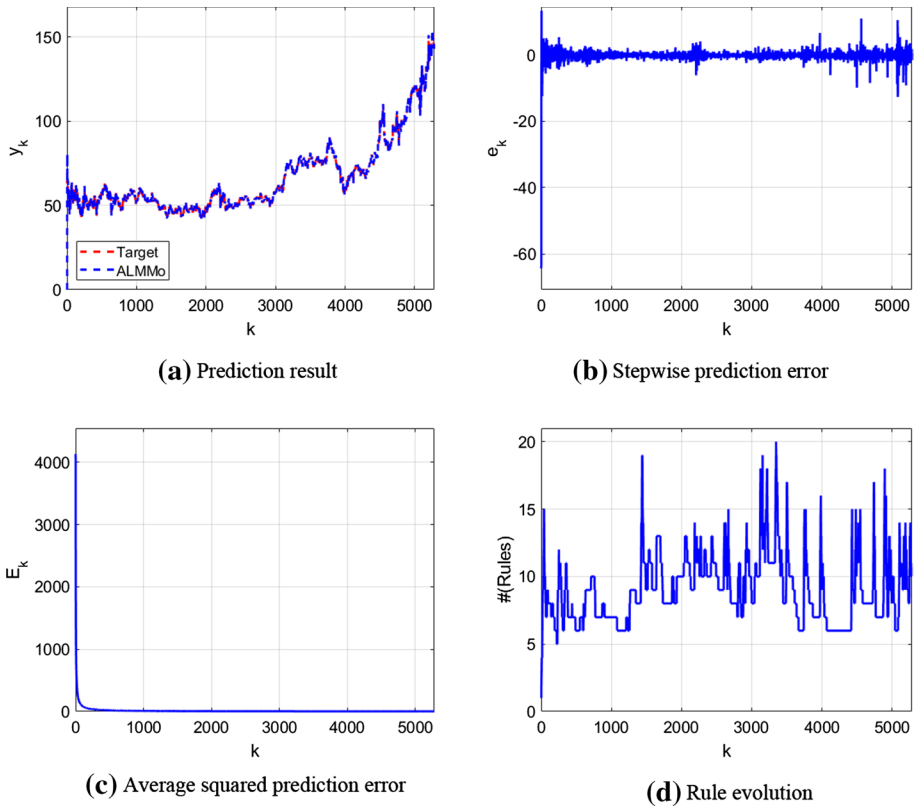
Fig. 5 Prediction result on WMT stock price

this example, the ALMMo system (Angelov et al. 2018) is employed for predicting the close price one day ahead based on the current four prices, namely,

$$\hat{x}_{k+1,close} = f\left(\left[x_{k,high}, x_{k,low}, x_{k,open}, x_{k,close}\right]^T\right) \tag{19}$$

The prediction results, stepwise prediction error ($e_k = x_{k+1,close} - \hat{x}_{k+1,close}$), average squared prediction error over time ($E_k = \frac{1}{k}\sum_{i=1}^{k} e_i^2$) and #(Rules) over the online learning process are shown in Fig. 5a–d, respectively.

Figure 5 shows that ALMMo is able to capture the dynamically changing patterns of the stock price and accurately modelling this nonstationary data stream (Fig. 5a). Despite that the uncertain nature of the stock market makes it impossible to correctly predict the stock price at all time (see Fig. 5b), it can be observed from Fig. 5c that the average squared prediction error of ALMMo converges to a very small value close to zero quickly and becomes stable afterwards. This demonstrates that the prediction performance of ALMMo is stable over time thanks to its capability of self-evolving its system structure and parameters in real time to follow the concept drifts and/or shifts in the data streams (Fig. 5d).

To illustrate the interpretability of fuzzy systems, the complete fuzzy rule base of the ALMMo system at the final time instance is tabulated in Table 10. The 10 first-order

Table 10 First-order fuzzy rules learned from historical WMT stock prices for prediction

# Rule	Detailed expression
$\mathbf{R}_1$	$IF\left(\begin{bmatrix} x_{k,high} \\ x_{k,low} \\ x_{k,open} \\ x_{k,close} \end{bmatrix} \sim \begin{bmatrix} 42.6800 \\ 42.0900 \\ 42.5900 \\ 42.2700 \end{bmatrix}\right)$ $THEN\left(y_1 = 0.0268 + 0.2525x_{k,high} + 0.2743x_{k,low} - 0.1997x_{k,open} + 0.6730x_{k,close}\right)$
$\mathbf{R}_2$	$IF\left(\begin{bmatrix} x_{k,high} \\ x_{k,low} \\ x_{k,open} \\ x_{k,close} \end{bmatrix} \sim \begin{bmatrix} 54.4900 \\ 49.7950 \\ 53.5950 \\ 51.9100 \end{bmatrix}\right)$ $THEN\left(y_2 = 0.0212 + 0.2371x_{k,high} + 0.2334x_{k,low} - 0.1568x_{k,open} + 0.6864x_{k,close}\right)$
$\mathbf{R}_3$	$IF\left(\begin{bmatrix} x_{k,high} \\ x_{k,low} \\ x_{k,open} \\ x_{k,close} \end{bmatrix} \sim \begin{bmatrix} 52.4000 \\ 51.3200 \\ 51.8100 \\ 51.9300 \end{bmatrix}\right)$ $THEN\left(y_3 = 0.0257 + 0.2066x_{k,high} + 0.2370x_{k,low} - 0.1744x_{k,open} + 0.7310x_{k,close}\right)$
$\mathbf{R}_4$	$IF\left(\begin{bmatrix} x_{k,high} \\ x_{k,low} \\ x_{k,open} \\ x_{k,close} \end{bmatrix} \sim \begin{bmatrix} 59.2900 \\ 58.3200 \\ 58.6100 \\ 58.3500 \end{bmatrix}\right)$ $THEN\left(y_4 = 0.1296 + 0.2698x_{k,high} + 0.2899x_{k,low} - 0.2196x_{k,open} + 0.6593x_{k,close}\right)$
$\mathbf{R}_5$	$IF\left(\begin{bmatrix} x_{k,high} \\ x_{k,low} \\ x_{k,open} \\ x_{k,close} \end{bmatrix} \sim \begin{bmatrix} 127.0000 \\ 117.6300 \\ 124.4600 \\ 119.4500 \end{bmatrix}\right)$ $THEN\left(y_5 = 3.0520 + 0.0269x_{k,high} + 0.4989x_{k,low} - 0.3317x_{k,open} + 0.7889x_{k,close}\right)$
$\mathbf{R}_6$	$IF\left(\begin{bmatrix} x_{k,high} \\ x_{k,low} \\ x_{k,open} \\ x_{k,close} \end{bmatrix} \sim \begin{bmatrix} 133.6300 \\ 128.8900 \\ 131.3300 \\ 129.5200 \end{bmatrix}\right)$ $THEN\left(y_6 = 2.7419 + 0.1944x_{k,high} + 0.2507x_{k,low} - 0.2044x_{k,open} + 0.7412x_{k,close}\right)$
$\mathbf{R}_7$	$IF\left(\begin{bmatrix} x_{k,high} \\ x_{k,low} \\ x_{k,open} \\ x_{k,close} \end{bmatrix} \sim \begin{bmatrix} 145.5838 \\ 143.4746 \\ 144.6488 \\ 144.3319 \end{bmatrix}\right)$ $THEN\left(y_7 = 1.9051 + 0.1541x_{k,high} + 0.8810x_{k,low} - 0.2832x_{k,open} + 0.2395x_{k,close}\right)$
$\mathbf{R}_8$	$IF\left(\begin{bmatrix} x_{k,high} \\ x_{k,low} \\ x_{k,open} \\ x_{k,close} \end{bmatrix} \sim \begin{bmatrix} 151.9800 \\ 150.2700 \\ 151.7400 \\ 151.3600 \end{bmatrix}\right)$ $THEN\left(y_8 = 1.1208 - 0.3303x_{k,high} + 0.7920x_{k,low} - 0.0971x_{k,open} + 0.6335x_{k,close}\right)$
$\mathbf{R}_9$	$IF\left(\begin{bmatrix} x_{k,high} \\ x_{k,low} \\ x_{k,open} \\ x_{k,close} \end{bmatrix} \sim \begin{bmatrix} 152.7200 \\ 151.1000 \\ 152.1500 \\ 151.6000 \end{bmatrix}\right)$ $THEN\left(y_9 = 1.0968 - 0.1705x_{k,high} + 0.9166x_{k,low} - 0.2446x_{k,open} + 0.4968x_{k,close}\right)$

Table 10 (continued)

# Rule	Detailed expression

$\mathbf{R}_{10}$

$$IF\left(\begin{bmatrix} x_{k,high} \\ x_{k,low} \\ x_{k,open} \\ x_{k,close} \end{bmatrix} \sim \begin{bmatrix} 153.6600 \\ 151.6600 \\ 153.6000 \\ 152.6400 \end{bmatrix}\right)$$
$$THEN\left(y_{10} = 1.1650 + 0.0069x_{k,high} + 1.1061x_{k,low} - 0.3599x_{k,open} + 0.2436x_{k,close}\right)$$

Table 11 Performance comparison on stock price prediction

Model	WMT	AAPL	COKE	MSFT
ALMMo	0.3020	0.1621	0.1025	0.1213
PSO-ALMMo	0.2872	0.1628	0.1043	0.1349
SVM	1.0450	4.9128	1.9897	4.1565
RF	0.3143	1.1183	0.8298	1.7168
MLP	0.3884	0.4889	0.5548	1.0864

fuzzy rules listed in this table is the "core" of ALMMo for fuzzy inferencing. Note that, all the parameters are self-learned from data directly.

The predictive performance of ALMMo is then tested on out-of-sample data collected within the period of 01.01.2021 to 01.01.2022 (1 year), 252 samples in total. In this experiment, the ALMMo system trained based on historical data (the previous 20 years) is used for predicting the one-day ahead close price of the WMT stock based on the daily four prices in 2021. However, different from the previous example before, the structure and parameters of ALMMo will not be updated during out-of-sample testing. The overall prediction error of ALMMo on WMT out-of-sample data in terms of *NDEI* is reported in Table 11, and is further compared against the following four approaches:

(1) PSO-ALMMo (the evolutionary version of ALMMo) (Gu et al. 2021a);
(2) Support vector machine regressor (SVM) (Cristianini and Shawe-Taylor 2000);
(3) Radom forest regressor (RF) (Breiman 2001), and;
(4) Multi-layer perceptron regressor (MLP) (Hastie et al. 2009).

In this example, PSO-ALMMo follows the same experimental setting as (Gu et al. 2021a); SVM uses the Gaussian kernel and the kernel scalar is selected using a heuristic procedure; RF is composed of 40 decision trees and the maximum split of each tree is $K - 1$ (K is the number of training samples); MLP has two hidden layers with 40 neurons in each and is trained using the resilient backpropagation algorithm. The prediction errors of PSO-ALMMo, SVM, RF and MLP are tabulated in Table 11 for comparison. The prediction results by ALMMo, PSO-ALMMo, SVM, RF and MLP are depicted in Fig. 6a.

In addition, the following three stocks, namely, Apple Inc (AAPL), Coca-Cola Consolidated Inc (COKE) and Microsoft Corporation (MSFT) are involved in this example, and the same experiments are repeated to predict the one-day ahead close prices of the three stocks under the same protocol. The prediction performances of the five approaches are also reported in Table 11.
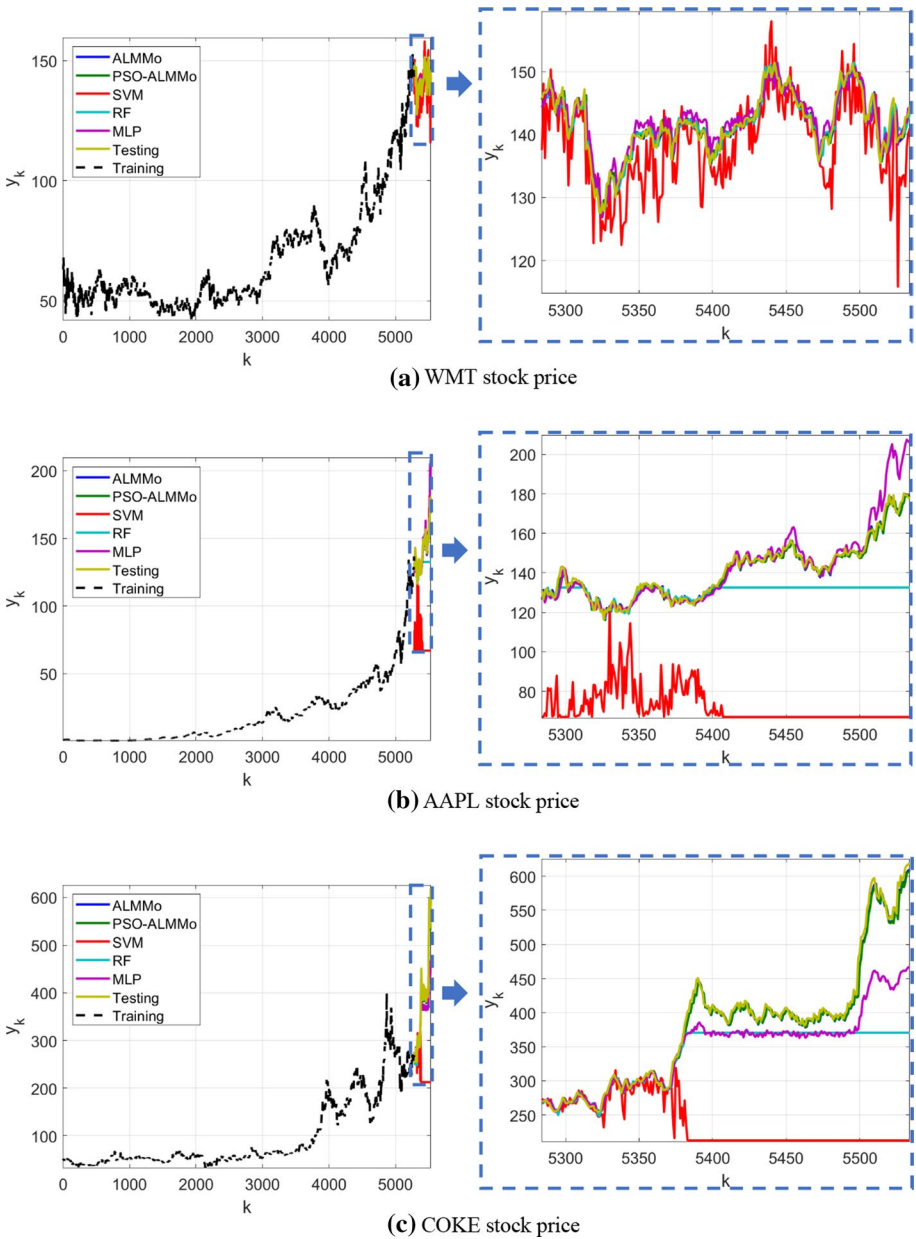
(a) WMT stock price

(b) AAPL stock price

(c) COKE stock price

Fig. 6 Prediction result comparison on four different stocks

It can be seen from Table 11 that the prediction performances of ALMMo and PSO-ALMMo are greater than the other three mainstream regressors, namely, SVM, RF and MLP. One can see clearly from Fig. 6 that the performances of SVM, RF and MLP deteriorate significantly when the stock prices exceed the value ranges of historical data used for training (see Fig. 6b–d), whilst both ALMMo and PSO-ALMMo are able to provide
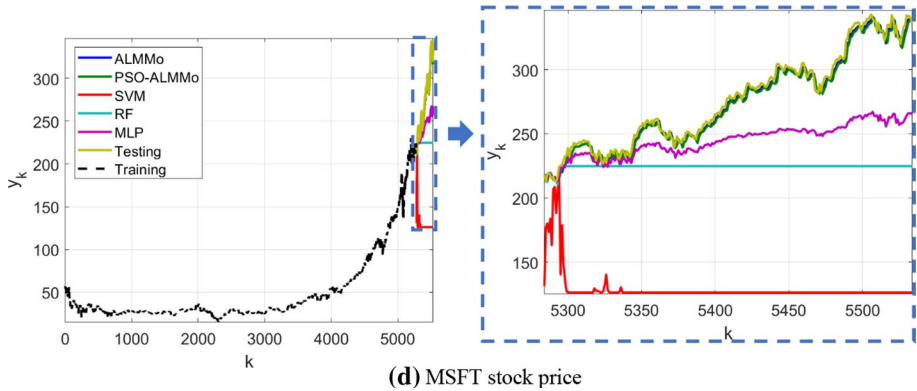
(d) MSFT stock price

Fig. 6 (continued)

accurate predictions consistently. For example, both SVM and RF failed to make reasonable predictions on MSFT after early 2021 when its stock price exceeded the historical highest point in the period of 2001–2020. The gap between the predictions made by MLP and the actual prices became increasingly larger as the stock price of MSFT keeps going higher. This performance comparison demonstrates the very strong capability of fuzzy systems in handling real-world uncertainties.

Another intrinsic advantage of fuzzy systems over alternative "black box" models, i.e., DNN and SVM, is their greater interpretability and transparency. The human-understandable, meaningful rule-based structure and parameters (see Table 10 for example) allow users to interpret the internal reasoning process and make sense of the predictions made by the model. This, in turn, makes the fuzzy rule-based models trustable for real-world applications.

9 Challenges and directions for further research and development

This paper has recalled the basic concepts of fuzzy sets and fuzzy systems and particularly, has introduced the main ideas of evolving, evolutionary and reinforcement learning-based fuzzy systems. It has provided an overview of the mainstream methods with an emphasis on their structure and parameter learning schemes. Then, a critical comparison between the three classes of fuzzy systems from the learning perspectives has been presented, enabling a better understanding of the pros and cons of the respective learning strategies. A high-level summary of real-world applications of fuzzy systems in different areas has been provided also.

To summarize, fuzzy systems have been widely recognized as a powerful tool for handling real-world problems with uncertainties by offering outstanding performance and high explainability. It has been a hotly researched field in recent decades, and there have been a wide variety of methods proposed to automate the design process of fuzzy systems. Many efforts have been devoted to further improve their learning performance, computational efficiency and model interpretability. Nevertheless, there remains a few open issues to be addressed as on-going and future work, including the following:

Explainability and transparency have increasingly become pressing issues in AI and machine learning due to the wider use of AI models in dealing with high-stakes and complex prediction applications in domains such as healthcare, finance, legal (Rudin 2019; Angelov et al. 2021). Although the state-of-the-art DNNs can offer greater performance on many challenging real-world problems (involving visual and speech information, in particular), such models are characterized as "black boxes" and can fool users (and even the systems developers themselves) easily. The lack of explainability and transparency can lead to severe or even fatal consequences (Rudin 2019). Although more complex models are not necessarily more accurate, it is generally recognised in practice that the simpler and more transparent fuzzy systems are incapable of handling high-dimensional, large-scale complex problems (Barredo Arrieta et al. 2020). To achieve greater performance in such application scenarios, the representation learning ability of fuzzy systems needs to be improved by redesigning the model structure, learning mechanism, or a mixture of both, and a trade-off between explainability/transparency and accuracy has to be made (Moral et al. 2021).

Missing data is commonly seen in many real-world applications and can have significantly adverse impact upon the conclusions drawn from data with missing values. Missing data may occur due to many different factors such as incomplete observations, transfer problems, memory loss, record damages, sensor failures, nonresponses, and so on (Škrjanc et al. 2019; Leite et al. 2020). There have been certain useful techniques for dealing with missing data, i.e., imputation (Musil et al. 2002), interpolation (Kokaram et al. 1995), and reverse engineering (Li et al. 2019), unsolved challenging issues exist, particularly in streaming data processing and multi-modal data modelling.

Class imbalance often occurs in real-world applications where the minorities are of greater interest, for example, in addressing problems such as financial fraud detection, network security, medical diagnosis and mechanical fault detection (Gu et al. 2020b; Naik et al. 2020). Conventional classifiers learned from imbalanced data tend to ignore the minorities because they are outnumbered and play a much weaker role in the overall performance evaluation. Although a number of successful methods (including a few evolutionary fuzzy systems (Sanz et al. 2015)) have been proposed for tackling this issue (López et al. 2013), learning an effective fuzzy rule-based model from imbalanced data streams is still a highly challenging task, especially when dealing with time-dependent or dynamic phenomena (Naik et al. 2018).

Curse of dimensionality refers to the various issues raised with high dimensional problems. Fuzzy systems are known to be less capable of handling high dimensional data as the systems can build a huge rule base from data, which usually causes significant overfitting effects and turns the systems into "black boxes" (Hagras 2018; Škrjanc et al. 2019). There exists a number of semantics-preserving dimensionality reduction or feature selection techniques, including those that are themselves based on fuzzy set theory, thereby easing the integration of such tools with the core fuzzy system (Jensen and Shen 2009). Evolutionary fuzzy systems are also capable of balancing the system complexity and performance by making a trade-off between different criteria using multi-objective evolutionary algorithms. In addition, learning efficiency may be gained through transformation of crisp rules while retaining model transparency (Chen et al. 2018). However, dimensionality reduction for high-dimensional data streams in the online application scenarios is fundamentally far more sophisticated considering the nonstationary nature of streaming data.

Big data is one of the hottest topics for current machine learning research. Single-model evolving fuzzy systems often fail to produce reliable results for large-scale problems, while evolutionary fuzzy systems are usually not applicable to such problems due to the very high computational complexity and additional system memory required. There have been

fuzzy rule-based systems proposed (López et al. 2014; Río et al. 2015) that use the MapReduce programming model (Dean and Ghemawat 2008), in an effort to learn and fuse fuzzy rule bases from big data. However, handling large-scale static data is not an easy task, and it is even more challenging to handle large-scale streaming data in real time. In such application scenarios, distributed or hierarchical ensemble learning frameworks would usually be a better option (Su et al. 2015; Gu et al. 2021b), but designing a suitable strategy for fuzzy systems without scarifying the interpretability is not straightforward.

Small data can also be an obstacle for machine learning algorithms. The vast majority of complex AI models benefit from the greater availability of labelled training data and require a lengthy computation procedure to learn the model. Although fuzzy systems are much less data hungry, building a precise, reliable, robust predictive modelling from few examples remains an immensely challenging task (Angelov and Gu 2018a, b). To date, there have been only very few fuzzy models that utilize semi-supervised learning techniques to involve unlabelled data during system identification (Gu and Angelov 2018b; Gu et al. 2022; Gu 2022). Even so, with limited data, it is very difficult, if not impossible, for the learned model to cover the entire problem space. To handle this important issue, one of the key development areas in the present fuzzy systems community is to establish fuzzy rule interpolation methods, enabling approximately inference to be performed (Naik et al. 2018; Li et al. 2021; Yang et al. 2022b). Nonetheless, the point of how to construct accurate fuzzy rule-based predictive models with extremely weak supervision is still to be addressed.

Ensemble learning is a powerful machine learning scheme to construct a stronger classifier by merging individual weaker classifiers (Polikar et al. 2001; Polikar 2006). The majority of existing works employ mainstream classifiers of other types, such as SVM (Xing and Liu 2020), MLP (Polikar et al. 2001), decision tree (Chen and Guestrin 2016), etc., and are designed for static data. There have been a few ensemble models that employ fuzzy systems as ensemble components to learn from (big) data streams (Scherer 2011; Soua et al. 2013; Iglesias et al. 2013b; Leite and Škrjanc 2019; Gu et al. 2021b; Lughofer et al. 2021; Lughofer and Pratama 2022), offering both great precision and high interpretability. Although the existing works on ensemble fuzzy models have reported promising results, only very few efforts have been made attempting to design novel ensemble frameworks specifically for better incorporating fuzzy systems (Gu and Angelov 2021). Hence, the potential of fuzzy rule-based systems in ensemble learning has not been fully explored. It would be worth investigating better ensemble strategies that make the best use of the human-interpretable fuzzy features offered by fuzzy systems. Another direction worth further exploration is the multi-layered deep ensemble fuzzy systems. It is well known that the power of DNNs comes from the multi-layered distributed representations. However, most of the existing ensemble fuzzy models are based on a flat structure, only very few works explore the possibility of constructing deep ensemble models with fuzzy systems (Pratama et al. 2020b; Gu 2021).

Finally, hybridization of fuzzy system and deep learning is a relatively new concept and has been hotly researched in the recent years. Despite of being criticized as "black box" models and fragile to uncertainties, DNNs have demonstrated impressive performances on various highly challenging image and natural language processing problems. Fuzzy systems can effectively handle uncertainties and offer greater interpretability and model-transparency, but could not reach the same levels of performance achieved by DNNs on these challenging tasks due to their simpler and smaller-scale internal structures. By integrating fuzzy systems with DNNs, it becomes possible to combine the advantages of both approaches. Currently, such hybridized approaches have been developed and applied

to image classification problems in the areas of remote sensing (Gu et al. 2022), autonomous driving (Soares et al. 2019a) and human activity recognition (Sargano et al. 2020), etc. These preliminary works have achieved promising performance, but alterative hybridization schemes to better integrate fuzzy systems and DNNs are still worth exploring in order to improve the performances and utility of the hybridized models. Further efforts are needed to implement these hybridized models for natural language processing problems, where DNNs are one of the dominant approaches used by researchers of this area.

References

Abadeh MS, Mohamadi H, Habibi J (2011) Design and analysis of genetic fuzzy systems for intrusion detection in computer networks. Expert Syst Appl 38(6):7067–7075

Adriaenssens V, De Baets B, Goethals PLM, De Pauw N (2004) Fuzzy rule-based models for decision support in ecosystem management. Sci Total Environ 319(1–3):1–12

Alcalá R, Alcalá-Fdez J, Herrera F (2007a) A proposal for the genetic lateral tuning of linguistic fuzzy systems and its interaction with rule selection. IEEE Trans Fuzzy Syst 15(4):616–635

Alcalá R, Gacto MJ, Herrera F, Alcalá-Fdez J (2007b) A multi-objective genetic algorithm for tuning and rule selection to obtain accurate and compact linguistic fuzzy rule-based systems. Int J Unc Fuzz Knowl Based Syst 15(5):539–557

Alcalá-Fdez J, Herrera F, Márquez F, Peregrín A (2007) Increasing fuzzy rules cooperation based on evolutionary adaptive inference systems. Int J Intell Syst 22(9):1035–1064

Alcalá-Fdez J et al (2011a) KEEL data-mining software tool: data set repository, integration of algorithms and experimental analysis framework. J Mult Log Soft Comput 17(2–3):255–287

Alcalá-Fdez J, Alcalá R, Herrera F (2011b) A fuzzy association rule-based classification model for high-dimensional problems with genetic rule selection and lateral tuning. IEEE Trans Fuzzy Syst 19(5):857–872

Alfaverh F, Denai M, Sun Y (2020) Demand response strategy based on reinforcement learning and fuzzy reasoning for home energy management. IEEE Access 8:39310–39321

Alghamdi M, Angelov P, Gimenez R, Rufino M, Soares E (2019) Self-organising and self-learning model for soybean yield prediction. Int Conf Soc Netw Anal Manag Secur 2019:441–446

Andonovski G, Angelov P, Blažič S, Škrjanc I (2016) A practical implementation of robust evolving cloud-based controller with normalized data space for heat-exchanger plant. Appl Soft Comput 48:29–38

Andonovski G, Lughofer E, Škrjanc I (2021b) Evolving fuzzy model identification of nonlinear Wiener-Hammerstein processes. IEEE Access 9:158470–158480

Andonovski G, Sipele O, Iglesias JA, Sanchis A, Lughofer E, Skrjanc I (2021a) Detection of driver maneuvers using evolving fuzzy cloud-based system. IEEE symposium series on computational intelligence. IEEE, pp 700–706

Andreu J, Angelov P (2010) Real-time human activity recognition from wireless sensors using evolving fuzzy systems. IEEE international conference on fuzzy systems. IEEE, pp 1–8

Angelov PP (1999) Evolving fuzzy rule-based models. International fuzzy systems association world congress. IEEE, pp 19–23

Angelov PP (2000) Evolving fuzzy rule-based models. J Chin Inst Ind Eng 17(5):459–468

Angelov PP (2002) Evolving rule-based models: a tool for design of flexible adaptive systems. Springer, Berlin Heidelberg

Angelov P, Buswell R (2002) Identification of evolving fuzzy rule-based models. IEEE Trans Fuzzy Syst 10(5):667–677

Angelov PP, Gu X (2018a) Toward anthropomorphic machine learning. IEEE Comput Mag 51(9):18–27

Angelov P, Guthke R (1997) A genetic-algorithm-based approach to optimization of bioprocesses described by fuzzy rules. Bioprocess Eng 16(5):299–303

Angelov PP, Yager R (2012) A new type of simplified fuzzy rule-based system. Int J Gen Syst 41(2):163–185

Angelov P, Zhou X (2008) Evolving fuzzy-rule based classifiers from data streams. IEEE Trans Fuzzy Syst 16(6):1462–1474

Angelov P (2010) Evolving Takagi-Sugeno fuzzy systems from streaming data. Evolving intelligent systems. John Wiley & Sons, USA, pp 21–50

Angelov P (2011) Fuzzily connected multimodel systems evolving autonomously from data streams. IEEE Trans Syst Man Cybern Part B Cybern 41(4):898–910

Angelov P (2012) Autonomous learning systems: from data streams to knowledge in real time. John Wiley & Sons Ltd

Angelov P, Buswell R (2001) Evolving rule-based models: a tool for intelligent adaption. IFSA world congress and 20th NAFIPS International conference. IEEE, pp 1062–1067

Angelov P, Filev D (2002) Flexible models with evolving structure. IEEE symposium on intelligent systems. IEEE, pp 28–33

Angelov P, Filev D (2003) On-line design of Takagi-Sugeno models. International fuzzy systems association world congress. Springer, pp 576–584

Angelov PP, Filev DP (2004) An approach to online identification of Takagi-Sugeno fuzzy model. IEEE Trans Syst Man Cybern Part B Cybern 34(1):484–498

Angelov P, Filev D (2005) Simpl_eTS: a simplified method for learning evolving Takagi-Sugeno fuzzy models. IEEE international conference on fuzzy systems. IEEE, pp 1068–1073

Angelov PP, Gu X (2018b) Empirical approach to machine learning. Springer International Publishing

Angelov PP, Filev DP, Kasabov NK (2010) Evolving intelligent systems: methodology and applications. Springer

Angelov PP, Gu X, Principe JC (2018) Autonomous learning multimodel systems from data streams. IEEE Trans Fuzzy Syst 26(4):2213

Angelov PP, Soares EA, Jiang R, Arnold NI, Atkinson PM (2021) Explainable artificial intelligence: an analytical review. Wires Data Min Knowl Discov 2021:1–13

Azad A et al (2020) Novel approaches for air temperature prediction: a comparison of four hybrid evolutionary fuzzy models. Meteorol Appl 27(1):1–12

Bao R, Rong H, Angelov PP, Chen B, Wong PK (2018) Correntropy-based evolving fuzzy neural system. IEEE Trans Fuzzy Syst 26(3):1324–1338

Barredo Arrieta A et al (2020) Explainable artificial antelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. Inf. Fusion 58:82–115

Berenji HR (1992) A reinforcement learning-based architecture for fuzzy logic control. Int J Approx Reason 6(2):267–292

Berenji HR, Khedkar P (1992) Learning and tuning fuzzy logic controllers through reinforcements. IEEE Trans Neural Netw 3(5):724–740

Blažič A, Škrjanc I, Logar V (2021) Soft sensor of bath temperature in an electric arc furnace based on a data-driven Takagi-Sugeno fuzzy model. Appl Soft Comput 113:107949

Blazic S, Skrjanc I (2020) Incremental fuzzy c-regression clustering from streaming data for local-model-network identification. IEEE Trans Fuzzy Syst 28(4):758–767

Breiman L (2001) Random forests. Mach Learn Proc 45(1):5–32

Camargos MO, Bessa I, D'Angelo MFSV, Cosme LB, Palhares RM (2020) Data-driven prognostics of rolling element bearings using a novel error based evolving Takagi-Sugeno fuzzy model. Appl Soft Comput J 96:106628

Cao B, Zhao J, Lv Z, Gu Y, Yang P, Halgamuge SK (2020) Multiobjective evolution of fuzzy rough neural network via distributed parallelism for stock prediction. IEEE Trans Fuzzy Syst 28(5):939–952

Carmona CJ, González P, del Jesus MJ, Navío-Acosta M, Jiménez-Trevino L (2011) Evolutionary fuzzy rule extraction for subgroup discovery in a psychiatric emergency department. Soft Comput 15(12):2435–2448

Casillas J, Cordón O, del Jesus MJ, Herrera F (2005) Genetic tuning of fuzzy rule deep structures preserving interpretability and its interaction with fuzzy rule set reduction. IEEE Trans Fuzzy Syst 13(1):13–29

Chang PC, Fan CY, Lin JJ (2011) Monthly electricity demand forecasting based on a weighted evolving fuzzy neural network approach. Int J Electr Power Energy Syst 33(1):17–27

Chatterjee A, Pulasinghe K, Watanabe K, Izumi K (2005) A particle-swarm-optimized fuzzy-neural network for voice-controlled robot systems. IEEE Trans Ind Electron 52(6):1478–1489

Chen T, Shang C, Su P, Shen Q (2018) Induction of accurate and interpretable fuzzy rules from preliminary crisp representation. Knowl Based Syst 146:152–166

Chen T, Guestrin C (2016) Xgboost: a scalable tree boosting system. ACM SIGKDD international conference on knowledge discovery and data mining. ACM, pp 785–794

Choi HH, Yun HM, Kim Y (2015) Implementation of evolutionary fuzzy PID speed controller for PM synchronous motor. IEEE Trans Ind Info 11(2):540–547

Cordón O (2011) A historical review of evolutionary learning methods for Mamdani-type fuzzy rule-based systems: designing interpretable genetic fuzzy systems. Int J Approx Reason 52(6):894–913

Cordón O, Herrera F (2000) A proposal for improving the accuracy of linguistic modeling. IEEE Trans Fuzzy Syst 8(3):335–344

Cordón O, Herrera F, Villar P (2001b) Generating the knowledge base of a fuzzy rule-based system by the genetic learning of the data base. IEEE Trans Fuzzy Syst 9(4):667–674

Cordón O, Herrera F, Magdalena L, Villar P (2001c) A genetic learning process for the scaling factors, granularity and contexts of the fuzzy rule-based system data base. Inf Sci 136(1–4):85–107

Cordón O, Herrera F, Hoffmann F, Magdalena L (2001a) Genetic fuzzy systems: evolutionary tuning and learning of fuzzy knowledge bases. World Scientific

Costa BSJ, Angelov PP, Guedes LA (2015) Fully unsupervised fault detection and identification based on recursive density estimation and self-evolving cloud-based classifier. Neurocomputing 150:289–303

Cristianini N, Shawe-Taylor J (2000) An introduction to support vector machines and other kernel-based learning methods. Cambridge University Press, Cambridge

Cunningham P, Delany SJ (2007) K-nearest neighbour classifiers. Mult Classif Syst 34:1–17

Dai X, Li CK, Rad AB (2005) An approach to tune fuzzy controllers based on reinforcement learning for autonomous vehicle control. IEEE Trans Intell Transp Syst 6(3):285–293

de Campos Souza PV (2020) Fuzzy neural networks and neuro-fuzzy networks: a review the main techniques and applications used in the literature. Appl Soft Comput 92:106275

de Jesús Rubio J (2009) SOFMLS: online self-organizing fuzzy modified least-squares network. IEEE Trans Fuzzy Syst 17(6):1296–1309

de Jesús Rubio J (2014) Evolving intelligent algorithms for the modelling of brain and eye signals. Appl Soft Comput 14:259–268

de Jesús Rubio J, Bouchachia A (2017) MSAFIS: an evolving fuzzy inference system. Soft Comput 21(9):2357–2366

de Souza PVC, Lughofer E (2020) Identification of heart sounds with an interpretable evolvin fuzzy neural network. Sensors 20(22):1–28

Dean J, Ghemawat S (2008) MapReduce: simplified data processing on large clusters. Commun ACM 51(1):107–113

Decker L, Leite D, Giommi L, Bonacorsi D (2020) Real-time anomaly detection in data centers for log-based predictive maintenance using an evolving fuzzy-rule-based approach. IEEE international conference on fuzzy systems. IEEE

del Río S, López V, Benítez JM, Herrera F (2015) A MapReduce approach to address big data classification problems based on the fusion of linguistic fuzzy rules. Int J Comput Intell Syst 8(3):422–437

Dosilovic FK, Brcic M, Hlupic N (2018) Explainable artificial intelligence: a survey. Int Conv Info Commun Technol Electron Microelectron MIPRO 2018:210–215

Dovžan D, Škrjanc I (2011) Recursive fuzzy c-means clustering for recursive fuzzy identification of time-varying processes. ISA Trans 50(2):159–169

Dovžan D, Logar V, Škrjanc I (2015) Implementation of an evolving fuzzy model (eFuMo) in a monitoring system for a waste-water treatment process. IEEE Trans Fuzzy Syst 23(5):1761–1776

Du W et al (2020) A new fuzzy logic classifier based on multiscale permutation entropy and its application in bearing fault diagnosis. Entropy 22(1):27

Duda RO, Hart PE, Stork DG (2000) Pattern classification. Wiley-Interscience, Chichester

Eberhart R, Kennedy J (1995) Particle swarm optimization. IEEE international conference on neural networks. IEEE, pp 1942–1948

Elhag S, Fernández A, Bawakid A, Alshomrani S, Herrera F (2015) On the combination of genetic fuzzy systems and pairwise learning for improving detection rates on intrusion detection systems. Expert Syst Appl 42(1):193–202

Elhag S, Fernández A, Alshomrani S, Herrera F (2019) Evolutionary fuzzy systems: a case study for intrusion detection systems. Evolutionary and swarm intelligence algorithms. Springer, Cham, pp 169–190

Esfahanipour A, Aghamiri W (2010) Adapted neuro-fuzzy inference system on indirect approach TSK fuzzy rule base for stock market analysis. Expert Syst Appl 37(7):4742–4748

Fahim M, Fatima I, Lee S, Park YT (2013) EFM: evolutionary fuzzy model for dynamic activities recognition using a smartphone accelerometer. Appl Intell 39(3):475–488

Fathinezhad F, Derhami V, Rezaeian M (2016) Supervised fuzzy reinforcement learning for robot navigation. Appl Soft Comput 40:33–41

Fazzolari M, Alcala R, Nojima Y, Ishibuchi H, Herrera F (2013) A review of the application of multi-objective evolutionary fuzzy systems: current status and further directions. IEEE Trans Fuzzy Syst 21(1):45–65

Fazzolari M, Alcalá R, Herrera F (2014) A multi-objective evolutionary method for learning granularities based on fuzzy discretization to improve the accuracy-complexity trade-off of fuzzy rule-based classification systems: D-MOFARC algorithm. Appl Soft Comput J 24:470–481

Feng HM, Chou HC (2012) Evolutionary fuzzy stock prediction system design and its application to the Taiwan stock index. Int J Innov Comput Inf Control 8(9):6173–6190

Feng J, Yu Y, Zhou ZH (2018) Multi-layered gradient boosting decision trees. Advances in neural information processing systems. IEEE, pp 3551–3561

Ferdaus MM, Pratama M, Anavatti SG, Garratt MA (2019) PALM: an incremental construction of hyperplanes for data stream regression. IEEE Trans Fuzzy Syst 27(11):2115–2129

Fernández A, López V, Del Jesus MJ, Herrera F (2015) Revisiting evolutionary fuzzy systems: taxonomy, applications, new trends and challenges. Knowl Based Syst 80:109–121

Fernandez A, Herrera F, Cordon O, Jose Del Jesus M, Marcelloni F (2019) Evolutionary fuzzy systems for explainable artificial intelligence: why, when, what for, and where to? IEEE Comput Intell Mag 14(1):69–81

Gacto MJ, Alcalá R, Herrera F (2009) Adaptation and application of multi-objective evolutionary algorithms for rule reduction and parameter tuning of fuzzy rule-based systems. Soft Comput 13(5):419–436

Gacto MJ, Alcalá R, Herrera F (2011) Interpretability of linguistic fuzzy rule-based systems: an overview of interpretability measures. Inf Sci 181(20):4340–4360

Garcia C, Esmin A, Leite D, Škrjanc I (2019) Evolvable fuzzy systems from data streams with missing values: with application to temporal pattern recognition and cryptocurrency prediction. Pattern Recognit Lett 128:278–282

Garibaldi JM (2019) The need for fuzzy AI. IEEE/CAA J Autom Sin 6(3):610–622

Ge D, Zeng XJ (2018a) Learning evolving T-S fuzzy systems with both local and global accuracy—a local online optimization approach. Appl Soft Comput 86:795–810

Ge D, Zeng XJ (2018b) A self-evolving fuzzy system which learns dynamic threshold parameter by itself. IEEE Trans Fuzzy Syst 27(8):1625–1637

Ge D, Zeng XJ (2020) Learning data streams online - an evolving fuzzy system approach with self-learning/adaptive thresholds. Inf Sci 507:172–184

Goharimanesh M, Mehrkish A, Janabi-Sharifi F (2020) A fuzzy reinforcement learning approach for continuum robot control. J Intell Robot Syst Theory Appl 100(3–4):809–826

Gorzałczany MB, Rudziński F (2017) Interpretable and accurate medical data classification—a multi-objective genetic-fuzzy optimization approach. Expert Syst Appl 71:26–39

Gu X (2021) Multilayer ensemble evolving fuzzy inference system. IEEE Trans Fuzzy Syst 29(8):2425–2431

Gu X (2022) An explainable semi-supervised self-organizing fuzzy inference system for streaming data classification. Inf Sci 583:364–385

Gu X, Angelov PP (2018a) Self-organising fuzzy logic classifier. Inf Sci 447:36–51

Gu X, Angelov PP (2018b) Semi-supervised deep rule-based approach for image classification. Appl Soft Comput 68:53–68

Gu X, Angelov P (2021) Multi-class fuzzily weighted adaptive boosting-based self-organising fuzzy inference ensemble systems for classification. IEEE Trans Fuzzy Syst. https://doi.org/10.1109/TFUZZ.2021.3126116

Gu X, Shen Q (2021) A self-adaptive fuzzy learning system for streaming data prediction. Inf Sci 579:623–647

Gu X, Angelov PP, Zhang C, Atkinson PM (2018) A massively parallel deep rule-based ensemble classifier for remote sensing scenes. IEEE Geosci Remote Sens Lett 15(3):345–349

Gu X, Shen Q, Angelov P (2021a) Particle swarm optimized autonomous learning fuzzy system. IEEE Trans Cybern 51(11):5352–5363

Gu X, Angelov P, Zhao Z (2021b) Self-organizing fuzzy inference ensemble system for big streaming data classification. Knowl Based Syst 218:106870

Gu X, Angelov P, Zhang C, Atkinson PM (2022) A semi-supervised deep rule-based approach for complex satellite sensor image analysis. IEEE Trans Pattern Anal Mach Intell 44(5):2281–2292

Gu X, Angelov PP, Ali AM, Gruver WA, Gaydadjiev G (2016) Online evolving fuzzy rule-based prediction model for high frequency trading financial data stream. IEEE conference on evolving and adaptive intelligent systems (EAIS). IEEE, pp 169–175

Gu S, Holly E, Lillicrap T, Levine S (2017) Deep reinforcement learning for robotic manipulation with asynchronous off-policy updates. IEEE international conference on robotics and automation. IEEE, pp 3389–3396

Gu X, Khan MA, Angelov P, Tiwary B, Yourdshah ES, Yang ZX (2020a) A novel self-organizing PID approach for controlling mobile robot locomotion. IEEE international conference on fuzzy systems. IEEE, pp 1–10

Gu X, Angelov PP, Soares EA (2020b) A self-adaptive synthetic over-sampling technique for imbalanced classification. Int J Intell Syst 35:6

Hadavandi E, Shavandi H, Ghanbari A (2010) A genetic fuzzy expert system for stock price forecasting. International conference on fuzzy systems and knowledge discovery. IEEE, pp 41–44

Hagras H (2018) Toward human-understandable, explainable AI. Computer 51(9):28–36

Han HG, Zhang L, Liu HX, Qiao JF (2018) Multiobjective design of fuzzy neural network controller for wastewater treatment process. Appl Soft Comput J 67:467–478

Hastie T, Tibshirani R, Friedman J (2009) The elements of statistical learning: data mining, inference, and prediction. Springer, Burlin

Hein D, Hentschel A, Runkler T, Udluft S (2017) Particle swarm optimization for generating interpretable fuzzy reinforcement learning policies. Eng Appl Artif Intell 65(August):87–98

Herrera F (2008) Genetic fuzzy systems: taxonomy, current research trends and prospects. Evol Intell 1(1):27–46

Hinojosa VH, Hoese A (2010) Short-term load forecasting using fuzzy inductive reasoning and evolutionary algorithms. IEEE Trans Power Syst 25(1):565–574

Holland JH (1975) Adaptation in natural and artificial systems. University of Michigan Press, Ann Arbor

Homaifar A, McCormick E (1995) Simultaneous design of membership functions and rule sets for fuzzy controllers using genetic algorithms. IEEE Trans Fuzzy Syst 3(2):129–139

Huang CJ, Lai WK, Hsiao SY, Liu HY, Luo RL (2004) A Bluetooth routing protocol using evolving fuzzy neural networks. Int J Wirel Inf Netw 11(3):131–146

Huang H, Rong HJ, Yang ZX, Vong CM (2021) Jointly evolving and compressing fuzzy system for feature reduction and classification. Inf Sci 579:218–230

Huang J, Angelov PP, Yin C (2020) Interpretable policies for reinforcement learning by empirical fuzzy sets. Eng Appl Artif Intell 91:103559

Iglesias JA, Angelov P, Ledezma A, Sanchis A (2010) Human activity recognition based on evolving fuzzy systems. Int J Neural Syst 20(5):355–364

Iglesias JA, Angelov P, Ledezma A, Sanchis A (2012) Creating evolving user behavior profiles automatically. IEEE Trans Knowl Data Eng 24(5):854–867

Iglesias JA, Ledezma A, Sanchis A (2013a) Evolving systems for computer user behavior classification. IEEE conference on evolving and adaptive intelligent systems. IEEE, pp 78–83

Iglesias JA, Ledezma A, Sanchis A (2013b) Ensemble method based on individual evolving classifiers. IEEE conference on evolving and adaptive intelligent systems. IEEE, pp 56–61

Ishibuchi H, Nozaki K, Yamamoto N, Tanaka H (1995) Selecting fuzzy if-then rules for classification problems using genetic algorithms. IEEE Trans Fuzzy Syst 3(3):260–270

Ishibuchi H, Murata T, Türkşen IB (1997) Single-objective and two-objective genetic algorithms for selecting linguistic rules for pattern classification problems. Fuzzy Sets Syst 89(2):135–150

Ishibuchi H, Nakashima T, Murata T (1999) Performance evaluation of fuzzy classifier systems for multidimensional pattern classification problems. IEEE Trans Syst Man Cybern Part B Cybern 29(5):601–618

Jensen R, Shen Q (2009) New approaches to fuzzy-rough feature selection. IEEE Trans Fuzzy Syst 17(4):824–838

Jouffe L (1998) Fuzzy inference system learning by reinforcement methods. IEEE Trans Syst Man Cybern C Appl Rev 28(3):338–355

Juang CF, Lu CH, Huang CA (2020) Navigation of three cooperative object-transportation robots using a multistage evolutionary fuzzy control approach. IEEE Trans Cybern. https://doi.org/10.1109/TCYB.2020.3015960

Juang CF, Lu CM (2009) Ant colony optimization incorporated with fuzzy Q-learning for reinforcement fuzzy control. IEEE Trans Syst Man Cybern Part A Syst Hum 39(3):597–608

Kadirkamanathan V, Niranjan M (1993) A function estimation approach to sequential learning with neural networks. Neural Comput 5(6):954–975

Karr C (1991) Genetic algorithms for fuzzy controllers. Artif Intell Expert 6(2):26–33

Kasabov N (2001) Evolving fuzzy neural networks for supervised/unsupervised online knowledge-based learning. IEEE Trans Syst Man Cybern Part B 31(6):902–918

Kasabov NK (2007) Evolving connectionist systems: the knowledge engineering approach. Springer Science & Business Media

Kasabov NK, Song Q (2002) DENFIS: dynamic evolving neural-fuzzy inference system and its application for time-series prediction. IEEE Trans Fuzzy Syst 10(2):144–154

Khalid K, Woungang I, Dhurandher SK, Singh J (2021) Reinforcement learning-based fuzzy geocast routing protocol for opportunistic networks. Internet Things 14:100384

Kim D, Choi YS, Lee SY (2002) An accurate COG defuzzifier design using Lamarckian co-adaptation of learning and evolution. Fuzzy Sets Syst 130(2):207–225

Kiran BR et al (2021) Deep reinforcement learning for autonomous driving: a survey. IEEE Trans Intell Transp Syst. https://doi.org/10.1109/TITS.2021.3054625

Kohonen T (1995) Learning vector quantization. In: Maps SO (ed) Berlin. Springer, Heidelberg, pp 175–189

Kokaram AC, Morris RD, Fitzgerald WJ, Rayner PJW (1995) Interpolation of missing data in image sequences. IEEE Trans Image Process 4(11):1509–1519

Konda VR, Tsitsiklis JN (2000) Actor-critic algorithms. Advances in neural information processing systems. IEEE, pp 1008–1014

Kukker A, Sharma R (2021) A genetic algorithm assisted fuzzy Q-learning epileptic seizure classifier. Comput Electr Eng 92:107154

LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nat Methods 13(1):35–35

Lee CC (1990) Fuzzy logic in control systems: fuzzy logic controller—Part 1&2. IEEE Trans Syst Man Cybern 20(2):404–435

Leite D, Škrjanc I (2019) Ensemble of evolving optimal granular experts, OWA aggregation, and time series prediction. Inf Sci 504:95–112

Leite D, Palhares RM, Campos VCS, Gomide F (2015) Evolving granular fuzzy model-based control of nonlinear dynamic systems. IEEE Trans Fuzzy Syst 23(4):923–938

Leite D, Škrjanc I, Gomide F (2020) An overview on evolving systems and learning from stream data. Evol Syst 11(2):181–198

Lekkas S, Mikhailov L (2010) Evolving fuzzy medical diagnosis of Pima Indians diabetes and of dermatological diseases. Artif Intell Med 50(2):117–126

Lemos A, Caminhas W, Gomide F (2011) Multivariable gaussian evolving fuzzy modeling system. IEEE Trans Fuzzy Syst 19(1):91–104

Lemos A, Caminhas W, Gomide F (2013) Adaptive fault detection and diagnosis using an evolving fuzzy classifier. Inf Sci 220:64–85

Leng G, Prasad G, McGinnity TM (2004) An on-line algorithm for creating self-organizing fuzzy neural networks. Neural Netw 17(10):1477–1493

Leng G, McGinnity TM, Prasad G (2005) An approach for on-line extraction of fuzzy rules using a self-organising fuzzy neural network. Fuzzy Sets Syst 150(2):211–243

Lewis FL, Vrabie D (2012) Reinforcement learning and feedback control: using natural decision methods to design optimal adaptive controllers. IEEE Control Syst 32(November):76–105

Li DZ, Wang W, Ismail F (2014) An evolving fuzzy neural predictor for multi-dimensional system state forecasting. Neurocomputing 145:381–391

Li Y, Zhang H, Xue X, Jiang Y, Shen Q (2018) Deep learning for remote sensing image classification: a survey. Wires Data Min Knowl Discov 8(6):e1264

Li P, Hastie TJ, Church KW (2006) Very sparse random projections. ACM SIGKDD international conference on knowledge discovery and data mining. ACM, pp 287–296

Li F, Li Y, Shang C, Shen Q (2019) Fuzzy knowledge-based prediction through weighted rule interpolation. IEEE Trans. Cybern. 2019:1–10

Li F, Shang C, Li Y, Yang J, Shen Q (2021) Approximate reasoning with fuzzy rule interpolation: background and recent advances. Springer

Lin CJ, Lin CT (1996) Reinforcement learning for an ART-based fuzzy adaptive learning control network. IEEE Trans Neural Netw 7(3):709–731

Lin FJ, Lin CH, Shen PH (2001) Self-constructing fuzzy neural network speed controller for permanent-magnet synchronous motor drive. IEEE Trans Fuzzy Syst 9(5):751–759

Lin CT, George Lee CS (1994) Reinforcement structure/parameter learning for neural-network-based fuzzy logic control systems. IEEE Trans Fuzzy Syst 2(1):46–63

Lin CT, Jou CP (2000) GA-based fuzzy reinforcement learning for control of a magnetic bearing system. IEEE Trans Syst Man Cybern Part B Cybern 30(2):276–289

Litjens G et al (2017) A survey on deep learning in medical image analysis. Med Image Anal 42:60–88

Liu YT, Lin YY, Wu SL, Chuang CH, Lin CT (2016) Brain dynamics in predicting driving fatigue using a recurrent self-evolving fuzzy neural network. IEEE Trans Neural Netw Learn Syst 27(2):347–360

López V, Fernández A, García S, Palade V, Herrera F (2013) An insight into classification with imbalanced data: empirical results and current trends on using data intrinsic characteristics. Inf Sci 250:113–141

López V, Del Río S, Benítez JM, Herrera F (2014) On the use of MapReduce to build linguistic fuzzy rule based classification systems for big data. IEEE international conference on fuzzy systems. IEEE, pp 1905–1912

Lughofer ED (2008) FLEXFIS: a robust incremental learning approach for evolving Takagi-Sugeno fuzzy models. IEEE Trans Fuzzy Syst 16(6):1393–1410

Lughofer E (2011a) Evolving fuzzy systems-methodologies, advanced concepts and applications. Springer, Berlin

Lughofer E (2011b) Evolving fuzzy systems—methodologies, advanced concepts and applications. Springer

Lughofer E (2021) Improving the robustness of recursive consequent parameters learning in evolving neuro-fuzzy systems. Inf Sci 545:555–574

Lughofer E, Angelov P (2011) Handling drifts and shifts in on-line data streams with evolving fuzzy systems. Appl Soft Comput 11(2):2057–2068

Lughofer E, Pratama M (2018) Online active learning in data stream regression using uncertainty sampling based on evolving generalized fuzzy models. IEEE Trans Fuzzy Syst 26(1):292–309

Lughofer E, Pratama M (2022) Online sequential ensembling of predictive fuzzy systems. Evol Syst 13(2):361–386

Lughofer E, Bouchot JL, Shaker A (2011) On-line elimination of local redundancies in evolving fuzzy systems. Evol Syst 2(3):165–187

Lughofer E, Cernuda C, Kindermann S, Pratama M (2015) Generalized smart evolving fuzzy systems. Evol Syst 6(4):269–292

Lughofer E, Pratama M, Skrjanc I (2018) Incremental rule splitting in generalized evolving fuzzy systems for autonomous drift compensation. IEEE Trans Fuzzy Syst 26(4):1854–1865

Lughofer E, Pratama M, Škrjanc I (2021) Online bagging of evolving fuzzy systems. Inf Sci 570:16–33

Maciel L, Ballini R, Gomide F (2017) An evolving possibilistic fuzzy modeling approach for value-at-risk estimation. Appl Soft Comput 60:820–830

Mackey MC, Glass L (1977) Oscillation and chaos in physiological control systems. Science 197(4300):287–289

Malik H, Sharma R, Mishra S (2020) Fuzzy reinforcement learning based intelligent classifier for power transformer faults. ISA Trans 101:390–398

Malik H, Yadav AK (2021) A novel hybrid approach based on relief algorithm and fuzzy reinforcement learning approach for predicting wind speed. Sustain. Energy Technol. Assess 43:100920

Mamdani EH, Assilian S (1975) An experiment in linguistic synthesis with a fuzzy logic controller. Int J Man Mach Stud 7(1):1–13

Márquez FA, Peregrín A, Herrera F (2007) Cooperative evolutionary learning of linguistic fuzzy rules and parametric aggregation connectors for Mamdani fuzzy systems. IEEE Trans Fuzzy Syst 15(6):1162–1178

Mehmanpazir F, Asadi S (2017) Development of an evolutionary fuzzy expert system for estimating future behavior of stock price. J Ind Eng Int 13(1):29–46

Mendes J, Souza F, Araujo R (2017) Online evolving fuzzy control design: an application to a CSTR plant. IEEE international conference on industrial informatics. IEEE, USA, pp 218–225

Mitra P, Murthy CA, Pal SK (2002) Unsupervised feature selection using feature similarity. IEEE Trans Pattern Anal Mach Intell 24(3):301–312

Mohamed S, Hameed IA (2018) A GA-based adaptive neuro-Fuzzy controller for greenhouse climate control system. Alexandria Eng J 57(2):773–779

Moral A, Castiello C, Magdalena L, Mencar C (2021) Explainable fuzzy systems. Springer International Publishing

Muñoz P, Barco R, De La Bandera I (2015) Load balancing and handover joint optimization in LTE networks using fuzzy logic and reinforcement learning. Comput Netw 76:112–125

Musil CM, Warner CB, Yobas PK, Jones SL (2002) A comparison of imputation techniques for handling Missing data. West J Nurs Res 24(7):815–829

Musilek P, Kromer P, Rodway J, Prauzek M (2015) Pressure-based forecasting of next-day solar energy availability using evolutionary fuzzy rules. IEEE international conference on fuzzy systems. IEEE

Mylonas SK, Stavrakoudis DG, Theocharis JB (2013) GeneSIS: a GA-based fuzzy segmentation algorithm for remote sensing images. Knowl Based Syst 54:86–102

Naik N, Diao R, Shen Q (2018) Dynamic fuzzy rule interpolation and its application to intrusion detection. IEEE Trans Fuzzy Syst 26(4):1878–1892

Naik N, Shang C, Jenkins P, Shen Q (2020) D-FRI-Honeypot: a secure sting operation for hacking the hackers using dynamic fuzzy rule interpolation. IEEE Trans Emerg Top Comput Intell. https://doi.org/10.1109/TETCI.2020.3023447

Nguyen NN, Zhou WJ, Quek C (2015) GSETSK: a generic self-evolving TSK fuzzy neural network with a novel Hebbian-based rule reduction approach. Appl Soft Comput J 35:29–42

Novitasari DCR, Wahyuni D, Munir M, Hidayati I, Amin FM, Oktafianto K (2019) Automatic detection of breast cancer in mammographic image using the histogram oriented gradient (HOG) descriptor and deep rule based (DRB) classifier method. International conference on advanced mechatronics intelligent manufacture and industrial automation. IEEE, pp 185–190

Onieva E, Hernandez-Jayo U, Osaba E, Perallos A, Zhang X (2015) A multi-objective evolutionary algorithm for the tuning of fuzzy rule bases for uncoordinated intersections in autonomous driving. Inf Sci 321:14–30

Park D, Kandel A, Langholz G (1994) Genetic-based new fuzzy reasoning models with application to fuzzy control. IEEE Trans Syst Man Cybern 24(1):39–47

Pe CA, Sipper M (1999) A fuzzy-genetic approach to breast cancer diagnosis. Artif Intell Med 17(2):131–155

Pham DT, Karaboga D (1991) Optimum design of fuzzy logic controllers using genetic algorithms. J Syst Eng 1(2):114–118

Polikar R (2006) Ensemble based systems in decision making. IEEE Circuits Syst Mag 6(3):21–44

Polikar R, Udpa L, Member S, Udpa SS, Honavar V (2001) Learn++ : an incremental learning algorithm for supervised neural networks. IEEE Trans Syst Man Cybern Part C (Appl Rev) 31(4):497–508

Pratama M, Anavatti SG, Lughofer E (2014a) Genefis: toward an effective localist network. IEEE Trans Fuzzy Syst 22(3):547–562

Pratama M, Anavatti SG, Angelov PP, Lughofer E (2014b) PANFIS: a novel incremental learning machine. IEEE Trans Neural Netw Learn Syst 25(1):55–68

Pratama M, Anavatti SG, Er MJ, Lughofer ED (2015) pClass: an effective classifier for streaming examples. IEEE Trans Fuzzy Syst 23(2):369–386

Pratama M, Dimla E, Tjahjowidodo T, Pedrycz W, Lughofer E (2020a) Online tool condition monitoring based on parsimonious ensemble+. IEEE Trans Cybern 50(2):664–677

Pratama M, Pedrycz W, Webb GI (2020b) An incremental construction of deep neuro fuzzy system for continual learning of nonstationary data streams. IEEE Trans Fuzzy Syst 28(7):1315–1328

Prauzek M, Krömer P, Rodway J, Musilek P (2016) Differential evolution of fuzzy controller for environmentally-powered wireless sensors. Appl Soft Comput J 48:193–206

Precup RE, Filip HI, Rədac MB, Petriu EM, Preitl S, Dragoş CA (2014) Online identification of evolving Takagi-Sugeno-Kang fuzzy models for crane systems. Appl Soft Comput J 24:1155–1163

Precup RE, Teban TA, Albu A, Borlea AB, Zamfirach IA, Petriu EM (2020) Evolving fuzzy models for prosthetic hand myoelectric-based control. IEEE Trans Instrum Meas 69(7):4625–4636

Pulkkinen P, Koivisto H (2008) Fuzzy classifier identification using decision tree and multiobjective evolutionary algorithms. Int J Approx Reason 48(2):526–543

Qaddoum K, Hines EL, Iliescu DD (2013) Yield prediction for tomato greenhouse using EFuNN. Int. Sch. Res. Not. 2013:1–9

Qiao JF, Hou Y, Zhang L, Han HG (2018) Adaptive fuzzy neural network control of wastewater treatment process with multiobjective operation. Neurocomputing 275:383–393

Quinlan JR (1986) Induction of decision trees. Mach Learn 1(1):81–106

Rajab S, Sharma V (2019) An interpretable neuro-fuzzy approach to stock price forecasting. Soft Comput 23(3):921–936

Razavi R, Klein S, Claussen H (2010) A fuzzy reinforcement learning approach for self-optimization of coverage in LTE networks. Bell Labs Tech J 15(3):153–175

Razavi-Far R, Davilu H, Palade V, Lucas C (2009) Model-based fault detection and isolation of a steam generator using neuro-fuzzy networks. Neurocomputing 72(13–15):2939–2951

Rey MI, Galende M, Fuente MJ, Sainz-Palmero GI (2017) Multi-objective based fuzzy rule based systems (FRBSs) for trade-off improvement in accuracy and interpretability: a rule relevance point of view. Knowledge-Based Syst 127:67–84

Rodrigues Júnior SE, de Oliveira Serra GL (2020) Intelligent forecasting of time series based on evolving distributed neuro-fuzzy network. Comput Intell 36(3):1394–1413

Rodríguez-Fdez I, Mucientes M, Bugarín A (2016a) FRULER: fuzzy rule learning through evolution for regression. Inf Sci 354:1–18

Rodríguez-Fdez I, Mucientes M, Bugarín A (2016b) S-FRULER: scalable fuzzy rule learning through evolution for regression. Knowl Based Syst 110:255–266

Rong HJ, Angelov P, Gu X, Bai J-M (2018) Stability of evolving fuzzy systems based on data clouds. IEEE Trans Fuzzy Syst 26(5):2774–2784

Rong H, Yang Z, Wong PK (2019) Robust and noise-insensitive recursive maximum correntropy-based evolving fuzzy system. IEEE Trans Fuzzy Syst 28(9):2277–2284

Rong HJ, Sundararajan N, Bin Huang G, Saratchandran P (2006) Sequential adaptive fuzzy inference system (SAFIS) for nonlinear system identification and prediction. Fuzzy Sets Syst 157(9):1260–1275

Rong HJ, Sundararajan N, Bin Huang G, Zhao GS (2011) Extended sequential adaptive fuzzy inference system for classification problems. Evol Syst 2(2):71–82

Rudin C (2019) Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat Mach Intell 1(5):206–215

Rudzi F (2016) A multi-objective genetic optimization of interpretability-oriented fuzzy rule-based classifiers. Appl Soft Comput 38:118–133

Sadeghi-Tehran P, Angelov P (2011) Online self-evolving fuzzy controller for autonomous mobile robots. IEEE workshop on evolving and adaptive intelligent systems. IEEE, pp 100–107

Safari S, Ardehali MM, Sirizi MJ (2013) Particle swarm optimization based fuzzy logic controller for autonomous green power energy system with hydrogen storage. Energy Convers Manag 65:41–49

Samanta S, Pratama M, Sundaram S (2019) A novel spatio-temporal fuzzy inference system (SPATFIS) and its stability analysis. Inf Sci 505:84–99

Sanz JA, Bernardo D, Herrera F, Bustince H, Hagras H (2015) A compact evolutionary interval-valued fuzzy rule-based classification system for the modeling and prediction of real-world financial applications with imbalanced data. IEEE Trans Fuzzy Syst 23(4):973–990

Saputro A, Rahma A, Nhita F (2016) An implementation of local regression smoothing on evolving fuzzy algorithm for planting calendar forecasting based on rainfall. International conference on soft computing and data mining. IEEE, pp 152–161

Sargano AB, Gu X, Angelov P, Habib Z (2020) Human action recognition using deep rule-based classifier. Multimed Tools Appl 79:30653–30667

Saw JG, Yang MCK, Mo TSEC (1984) Chebyshev inequality with estimated mean and variance. Am Stat 38(2):130–132

Scherer R (2011) An ensemble of logical-type neuro-fuzzy systems. Expert Syst Appl 38(10):13115–13120

Serdio F, Lughofer E, Pichler K, Buchegger T, Efendic H (2014) Residual-based fault detection using soft computing techniques for condition monitoring at rolling mills. Inf Sci 259:304–320

Severiano CA, de Silva PCL, Weiss Cohen M, Guimarães FG (2021) Evolving fuzzy time series for spatio-temporal forecasting in renewable energy systems. Renew Energy 171:764–783

Shi Y, Eberhart R, Chen Y (1999) Implementation of evolutionary fuzzy systems. IEEE Trans Fuzzy Syst 7(2):109–119

Silva S, Costa P, Santana M, Leite D (2020) Evolving neuro-fuzzy network for real-time high impedance fault detection and classification. Neural Comput Appl 32(12):7597–7610

Silver D et al (2016) Mastering the game of Go with deep neural networks and tree search. Nature 529(7587):484–489

Škrjanc I (2021) An evolving concept in the identification of an interval fuzzy model of Wiener-Hammerstein nonlinear dynamic systems. Inf Sci 581:73–87

Škrjanc I, Andonovski G, Ledezma A, Sipele O, Iglesias JA, Sanchis A (2018a) Evolving cloud-based system for the recognition of drivers' actions. Expert Syst Appl 99:231–238

Škrjanc I, Ozawa S, Ban T, Dovžan D (2018b) Large-scale cyber attacks monitoring using evolving Cauchy possibilistic clustering. Appl Soft Comput 62:592–601

Škrjanc I, Iglesias J, Sanchis A, Leite D, Lughofer E, Gomide F (2019) Evolving fuzzy and neuro-fuzzy approaches in clustering, regression, identification, and classification: a survey. Inf Sci 490:344–368

Škrjanc I, Blažič S, Angelov P (2014) Robust evolving cloud-based PID control adjusted by gradient learning method. IEEE conference on evolving and adaptive intelligent systems. IEEE, pp 1–8

Soares E, Costa P, Costa B, Leite D (2018a) Ensemble of evolving data clouds and fuzzy models for weather time series prediction. Appl Soft Comput 64:445–453

Soares EA, Camargo HA, Camargo SJ, Leite DF (2018b) Incremental Gaussian granular fuzzy modeling applied to hurricane track forecasting. IEEE International conference on fuzzy systems. IEEE, pp 1–8

Soares E, Angelov P, Costa B, Castro M (2019a) Actively semi-supervised deep rule-based classifier applied to adverse driving scenarios. International joint conference on neural networks. IEEE, pp 1–8

Soares E, Angelov P, Filev D, Costa B, Castro M, Nageshrao S (2019b) Explainable density-based approach for self-driving actions classification. IEEE international conference on machine learning and applications. IEEE, pp 469–474

Soares E, Angelov P, Gu X (2020) Autonomous learning multiple-model zero-order classifier for heart sound classification. Appl. Soft Comput. J. 94:1–9

Soleimani-B H, Lucas C, Araabi BN (2010) Recursive Gath-Geva clustering as a basis for evolving neuro-fuzzy modeling. Evol Syst 1(1):59–71

Soua B, Borgi A, Tagina M (2013) An ensemble method for fuzzy rule-based classification systems. Knowl Inf Syst 36(2):385–410

Stavrakoudis DG, Theocharis JB, Zalidis GC (2011) A boosted genetic guzzy classifier for land cover classification of remote sensing imagery. ISPRS J Photogramm Remote Sens 66(4):529–544

Stirling J, Chen T, Bucholc M (2020) Diagnosing Alzheimer's disease using a self-organising fuzzy classifier. Fuzzy logic: recent applications and developments. IEEE, p 69

Su P, Shang C, Shen Q (2015) A hierarchical fuzzy cluster ensemble approach and its application to big data clustering. J Intell Fuzzy Syst 28(6):2409–2421

Subramanian K, Suresh S (2012) A meta-cognitive sequential learning algorithm for neuro-fuzzy inference system. Appl Soft Comput 12(11):3603–3614

Sugeno M, Kang GT (1988) Structure identification of fuzzy model. Fuzzy Sets Syst 28(1):15–33

Sutton RS, Barto AG (2018) Reinforcement learning: an introduction. MIT press

Takagi T, Sugeno M (1985) Fuzzy identification of systems and its applications to modeling and control. IEEE Trans Syst Man Cybern 15(1):116–132

Thrift P (1991) Fuzzy logic synthesis with genetic algorithms. International conference on genetic algorithms. Academia, pp 509–513

Tjoa E, Guan C (2020) A survey on explainable artificial intelligence (xai): toward medical xai. IEEE Trans Neural Netw Learn Syst. https://doi.org/10.1109/TNNLS.2020.3027314

Tsang CH, Kwong S, Wang H (2007) Genetic-fuzzy rule mining approach and evaluation of feature selection techniques for anomaly intrusion detection. Pattern Recognit 40(9):2373–2391

Tsourveloudis NC (2010) On the evolutionary-fuzzy control of WIP in manufacturing systems. Neurocomputing 73(4–6):648–654

Valenzuela-Rendon M (1991) The fuzzy classifier system: a classifier system for continuously varying variables. International conference on genetic algorithms. CiNii, pp 346–353

Wang XS, Cheng YH, Yi JQ (2007) A fuzzy actor-critic reinforcement learning network. Inf Sci 177(18):3764–3781

Wu Q et al (2021) A fuzzy-inference-based reinforcement learning method of overtaking decision making for automated vehicles. Proc Inst Mech Eng Part D J Automob Eng. 2021:09544070211018099

Xing HJ, Liu WT (2020) Robust AdaBoost based ensemble of one-class support vector machines. Inf Fusion 55:45–58

Yang ZX, Rong HJ, Wong PK, Angelov P, Yang ZX, Wang H (2021a) Self-evolving data cloud-based PID-like controller for nonlinear uncertain systems. IEEE Trans Ind Electron 68(5):4508–4518

Yang H et al (2021b) Intelligent reflecting surface assisted anti-jamming communications: a fast reinforcement learningapproach. IEEE Trans Wirel Commun 20(3):1963–1974

Yang ZX, Rong HJ, Angelov PP, Yang ZX (2022a) Statistically evolving fuzzy inference system for non-Gaussian noises. IEEE Trans Fuzzy Syst 30(4):2649–2664

Yang J, Shang C, Li Y, Li F, Shen L, Shen Q (2022b) Constructing anfis with sparse data through group-based rule interpolation: an evolutionary approach. IEEE Trans Fuzzy Syst 30(4):893–907

Ye C, Yung NHC, Wang D (2003) A fuzzy controller with supervised learning assisted reinforcement learning algorithm for obstacle avoidance. IEEE Trans Syst Man Cybern Part B 33(1):17–27

Yong Y, Lee Y, Gu X, Angelov P, Ling Ngo DC, Yourdshahi ES (2018) Foreign currency exchange rate prediction using neuro-fuzzy systems. Procedia Comput Sci. 144:232–238

Yu L, Zhang YQ (2005) Evolutionary fuzzy neural networks for hybrid financial prediction. IEEE Trans Syst Man Cybern Part C (appl Rev) 35(2):244–249

Yu Z, Dexter A (2010) Online tuning of a supervisory fuzzy controller for low-energy building system using reinforcement learning. Control Eng Pract 18(5):532–539

Yung NHC, Ye C (1999) An intelligent mobile vehicle navigator based on fuzzy logic and reinforcement learning. IEEE Trans Syst Man Cybern Part B Cybern 29(2):314–321

Zadeh LA (1965) Fuzzy sets. Inf Control 8(3):338–353

Zadeh LA (1973) Outline of a new approach to the analysis of complex systems and decision processes. IEEE Trans Syst Man Cybern 1:28–44

Zhang H, Zhang K, Cai Y, Han J (2019b) Adaptive fuzzy fault-tolerant tracking control for partially unknown systems with actuator faults via integral reinforcement learning method. IEEE Trans Fuzzy Syst 27(10):1986–1998

Zhang Q, Filev D, Szwabowski S, Langari R (2019a) A real-time fuzzy learning algorithm for Markov chain and its application on prediction of vehicle speed. IEEE international conference on fuzzy systems. IEEE, pp 1–6

Zhou ZH, Feng J (2017) Deep forest: towards an alternative to deep neural networks. International joint conference on artificial intelligence. IEEE, pp 3553–3559

Zhou C, Meng Q (2003) Dynamic balance of a biped robot using fuzzy reinforcement learning agents. Fuzzy Sets Syst 134(1):169–187

Zimmermann HJ (2010) Fuzzy set theory. Wiley Interdiscip Rev Comput Stat 2(3):317–332

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.