Evolutionary Fuzzy Systems for Explainable Artificial Intelligence: Why, When, What for, and Where to?

Alberto Fernandez, Maria Jose del Jesus, *Member, IEEE,* Oscar Cordon, *Fellow Member, IEEE,* Francesco Marcelloni, *Member, IEEE,* and Francisco Herrera, *Senior Member, IEEE*

Abstract

Evolutionary fuzzy systems are one of the greatest advances within the area of computational intelligence. They consist of evolutionary algorithms applied to the design of fuzzy systems. Thanks to this hybridization, superb abilities are provided to fuzzy modeling in many different data science scenarios. This contribution is intended to comprise a position paper developing a comprehensive analysis of the evolutionary fuzzy systems research field. To this end, the "4 W" questions are posed and addressed with the aim of understanding the current context of this topic and its significance. Specifically, it will be pointed out why evolutionary fuzzy systems are important from an explainable point of view, when they began, what they are used for, and where the attention of researchers should be directed to in the near future in this area. They must play an important role for the emerging area of eXplainable Artificial Inteligence (XAI) learning from data.

Index Terms

Evolutionary fuzzy systems, Fuzzy rule-based systems, Evolutionary algorithms, Explainable artificial intelligence, Interpretability, Data science, Big data.

I. Why Evolutionary Fuzzy Systems?

"Knowledge itself is power" [1]. This simple sentence has led to a wave of continuous interest in developing models that are able to extract the maximum amount of information from data. In this context, the use of machine learning (ML) techniques allows stakeholders to obtain useful insights, predictions, and decisions from datasets of many different sources in an automatic fashion [2]. Current applications come with novel data characteristics as big dimension and non standard classification problems, and both researchers and practitioners actually aim to

A. Fernandez, O. Cordon, and F. Herrera are with the DaSCI Andalusian Institute of Data Science and Computational Intelligence, University of Granada, Granada, 18017, Spain. e-mail: {alberto,ocordon,herrera}@decsai.ugr.es.

M.J. del Jesus is with the DaSCI Andalusian Institute of Data Science and Computational Intelligence, Department of Computer Science, University of Jaen, Jaen, 23071, Spain. email: mjjesus@ujaen.es.

F. Marcelloni is with the Department of Information Engineering, University of Pisa, Pisa, 56126, Italy. email: francesco.marcelloni@unipi.it.

Manuscript received - -, -; revised - -, -.

understand how the models work. Therefore, a movement is being witnessed from traditional data mining towards a more profitable and challenging scenario known as data science [3]. It brings novel technologies, frameworks, methodologies, and skills that are designed to ease the management of the challenges related to big data applications [4].

Nevertheless, to be able to reach the deepest level when considering all the information available, the knowledge domain and the data analysis must have strong synergy. From a data viewpoint, one must be aware of the quality associated with it. Additionally, representation of the expert knowledge available on the tackled application domain must also be considered.

In this scenario, fuzzy set theory might be regarded as a valuable tool. Proposed by Lofti A. Zadeh in the mid 60s [5], a linguistic representation of numerical variables is possible by means of a membership degree being assigned to each of them, which could vary from 0 (full non-membership) to 1 (full membership). Therefore, models that use fuzzy sets as a tool for data handling provide some important advantages. In terms of semantics, the use of linguistic labels in the fuzzy model structure is a natural knowledge representation allowing for a direct human interaction [6]. In addition, from a learning perspective, translating the input features into fuzzy variables with fuzzy membership functions permits obtaining smoothed descriptive models that adapt well to data with a certain degree of uncertainty.

Based on the former observations, and without lacking generality, the focus of this contribution will be set on fuzzy rule-based systems (FRBSs) [7]. As their name suggests, FRBSs are composed of fuzzy IF-THEN rules where both antecedents and consequents usually contain fuzzy sets. The main components of any FRBS are the knowledge base (KB) and the inference engine module. The KB comprises all the fuzzy rules within a rule base (RB), and the definition of the fuzzy sets in the data base (DB). The inference engine includes a fuzzification interface, an inference system, and a defuzzification interface.

The learning procedure for FRBSs involves a search for the model that, based on the observations, best approximates a given performance metric [2]. In particular, finding the best approach is related to the so-called "empirical risk-loss" function, which depends on the data the user works with. In other words, any learning algorithm has the ultimate goal of optimizing a mapping function between inputs and outputs. In this sense, the clear advantages of evolutionary algorithms (EAs) [8] for developing this task must be highlighted.

EAs are a type of metaheuristic based optimization techniques that are based on a population of solutions. As the name suggests, EAs are concerned with biological evolution, so that each solution is encoded as an individual of the population. The search for the optimal values of the function that is being optimized (the fitness value) is carried out by means of the repeated application of operators such as reproduction, mutation, recombination and selection. Among the different implementations of EAs, the most popular is certainly Genetic Algorithms (GAs) [9], where solutions (also known as chromosomes) are encoded into strings of "genes", and evolve by means of recombination and/or mutation. EAs are not specifically designed as ML techniques. However, it is well-known that a learning task can be modeled as an optimization problem and thus effectively solved through evolution. EAs' powerful search in complex, ill-defined problem spaces has allowed for them to be successfully applied to a huge variety of ML and knowledge discovery tasks [10]. When the accuracy performance is set as the fitness function of this optimization process, the predictive quality of the system is expected to be boosted.

Considering the issues on predictive performance, the main goal traditionally pursued is to make the model matching reality, i.e. effectively summarizing the underlying data. Hence, this “quest for accuracy” is the most important requirement when selecting a solution, and has probably supported the current explosion of black-box ML approaches. This type of system is known for having an excellent ability to learn accurately from the input data. On the contrary, most of these accurate models are highly non-transparent, i.e. it is no clear what information in the input data makes them arrive at their decisions [11], [12]. Nowadays, there is an important need of safety, ethic, and with scientific understanding systems that provide the right to explanation where necessary [13]. They must be optimized not only for accuracy but also for other criteria as fairness or unbiasedness, privacy, reliability, robustness, causality and/or trust, among others. Most of these criteria often cannot be completely quantified, but if the system is *interpretable*, i.e. if it can *explain* its reasoning, it can be verified whether that reasoning is sound with respect to these auxiliary criteria.

There is little consensus on the definition of interpretability, explanability and some of the auxiliary criteria. In the ML context the *interpretability* is defined as the ability to explain or to present in understable terms to a human [14]. In recent years has emerged the broader concept *explainability* in what has been called *eXplanaible Artificial Intelligence (XAI)* that encompasses ML/AI systems for opening black box models, for improving the understanding of what the models have learned and/or for explaining individual predictions [11].

The design of this kind of systems includes different aspects to be taken into account:

- Designers and developers must have the chance of analyzing the generated model and discerning its meaning. i.e. to understand the structure of the system.
- End-users must use the system as a decision support, so that the model must explain the phenomena under study.

Interpretability is probably the traditional buzz-word, in ML context, and *explainability* in the broader context of AI (XAI). Both with equivalent meaning and behind the desideratum for accomplishing all the former issues. However, different sub-concepts can be distinguished associated with terms “interpretability” and ”explainability”. Among others, we should refer to “understandability”, ”intelligibility” or “comprehensibility” [12], [15].

- *Understandability* and *intelligibility* can be viewed as synonyms that are associated, in ML, to a functional understanding of the model. That is, to grasp how the model works [16], without trying to elucidate its inner workings o shed light on its internal representation [17]. For FRBSs, intelligibility is primarily associated with inference.
- *Comprensibility* was defined as the learning algorithm ability for encoding its model in such a way that it may be inspected and understood by humans [15], [18]. This definition narrows the focus to the model itself. It is based on the comprehensibility postulate argued by Michalski [19]: *”The results of computer induction should be symbolic descriptions of given entities, semantically and structurally similar to those a human expert might produce observing the same entities. Components of these descriptions should be comprehensible as single ‘chunks’ of information, directly interpretable in natural language, and should relate quantitative*

and qualitative concepts in an integrated fashion". In the FRBS context, the comprensibily is related with the fuzzy rules, and in general with the KB.

The final objective related to explainability is boosting the transparency in the solutions proposed for data science applications, and therefore the ability to trust the system output [20].

One straightforward way for combining two of the most important concepts in ML (accuracy and interpretability/explainability) in a natural way, to obtaining XAI models, is by means of the synergy between FRBSs and EAs. Specifically, the intrinsic understandability, comprehensibility, and explainability associated with FRBSs, and the potential of EAs as optimization technique for improving FRBSs, leads to what is known as evolutionary fuzzy systems (EFSs) [21].

In short, an EFS is a methodology that utilizes EAs for a "fine-setting" of the components of the FRBS, with the aim of achieving a more reliable solution in terms of any desired objective function, based on accuracy, interpretability, or a combination of both. The EA may be applied *a priori* during the building stage of the FRBS, leading to a "learning" stage, or rather to be considered *a posteriori* for a finer adjustment of the FRBS, which is known as a "tuning" stage. Different perspectives for the application of both approaches may be found in the specialized literature. To set up the fuzzy sets, the parameterization of the fuzzy membership functions, and the selection of the fuzzy rules, among others, should be noted. In addition, a very interesting extension uses multi-objective evolutionary algorithms (MOEAs) [22], which can consider several design criteria to be optimized concurrently, thus composing the so-called multi-objective evolutionary fuzzy systems (MOEFSs) [23].

In this contribution, a better insight on the topic of EFS is provided by posing the *4 W questions*. The first one has been addressed throughout this section, namely *Why are EFSs so significant in data science tasks?* The main reason has been discussed above: they have a great potential for obtaining a good trade-off between accurate and explainable models, which is extremely important in most of the current social and engineering applications.

The remaining 3 W questions are analyzed in depth to set the past, present and future scenarios. Whereas the historical review (past and present) is intended to set the current context of EFSs, the main novelty in this position paper is related to future insights. Specifically, special attention is paid to analyze the capabilities of EFSs, in particular those areas that could benefit the most from their application. Additionally, some novel scenarios, namely data science, big data and the development of explainable models, that EFS researchers should focus on for the near future are described. Finally, the paper is concluded by a thorough discussion on the need for XAI models in ML, which is intended to go beyond the old-fashioned dialectical exercise on the interpretability-accuracy trade-off, very extended in FRBSs.

To carry out these tasks, the remainder of this paper is organized as follows. First, Section II considers *When did EFSs begin?*, introducing the pioneer works on the topic and the current taxonomy for EFSs. Section III analyzes *What should EFSs be used for?*, presenting the good capabilities of EFSs, namely robustness and explainability, for solving current problems and applications. Section IV is devoted to establishing *Where are EFSs going to?*, considering some challenges for future work, and focusing in particular on the big data scenario due to its significance for current research. To summarize all thoughts considered in this position paper, Section V presents an overview of the need for XAI. Finally, Section VI points out some concluding remarks.

II. When Did Evolutionary Fuzzy Systems Begin? Past and Present

In order to understand the present and to be able to forecast the future of a specific topic, the history and state of the art must be acknowledged. In this regard, this section highlights the main progresses made in EFSs from their proposal to present. To this aim, the initial milestones on the topic and some well-known reviews are first introduced (Section II-A). Then, a complete taxonomy is presented to frame the EFS techniques in their different categories (Section II-B).

A. EFSs: Pioneering approaches

The pioneering work by Lofti A. Zadeh on fuzzy sets supposed a significant contribution in many research areas of control and modeling [5]. Specifically, the application of fuzzy set theory as a knowledge representation structure based on fuzzy rules began in the mid-1970s with the work of Mamdani [7]. Practically at the same time, the basis and theoretical studies on EAs were established, in particular regarding one of its branches, that of GAs [9]. Despite the early appearance of both approaches, only in the early 1990s the hybridization between these two computational intelligence techniques was introduced.

Throughout this section, four of the pioneering works in the field are discussed in no particular order. They adopted both tuning and learning methodologies of the components of the FRBS KB by means of GAs.

Karr was the first to investigate the genetic tuning of the DB for fuzzy controllers [24]. In his study, the complete definition of the DB was intended to be optimized, i.e. the best values of the parameters that define the fuzzy partitions were automatically determined. This was obtained thanks to the ease of encoding information on the GA chromosome. In particular, the evolution of the input and output fuzzy sets were jointly considered.

One of the first models of evolutionary learning of a linguistic RB was proposed by Valenzuela-Rendon in [25]. This proposal used a learning scheme where each solution (chromosome) represented a single rule over the entire RB. The first version of the methodology used a reward distribution scheme. Later, the original proposal was extended to allow reinforcement learning.

A different scheme for learning the RB was defined by Thrift in [26]. The author used a decision matrix to represent the RB, in that particular case study only with two dimensions. This matrix collected the linguistic labels of the output variable according to the corresponding input fuzzy labels. It considered an integer coding scheme, using “0” to represent a null value, thus allowing for the automatic learning of the optimal number of rules. The GA encoded different fuzzy rules, i.e. a whole RB definition, on a single chromosome.

Pham and Karaboga proposed a similar approach but using a fuzzy relation R instead of the classical crisp relation (decision table) [27]. The GA was used to modify the fuzzy relational matrix of a single-input / single-output fuzzy model. The chromosome was obtained by concatenating the $M \cdot N$ elements of R, where M and N were the number of linguistic terms associated with the input and output variables, respectively. The main difference with the work in [26] is that the elements of R were real numbers in [0,1] instead of integer values.

All these contributions were considered as milestones for the work on EFSs. From that point on, the interest of researchers on the topic has significantly increased. As a short description of developments throughout the history of EFSs, four surveys are considered. Ten years after the publication of the first approaches, the main achievements in the field were compiled in the 2001 monograph by Cordón, Herrera, Hoffmann, and Magdalena [21]. After a period of another ten years, Cordón proposed a historical review focusing on the interpretability viewpoint in 2011 [28]. In 2013, Fazzolari, Alcalá, Nojima, Ishibuchi, and Herrera published an overview focused on the MOEFSs topic, which was intended to summarize the main contributions in this particular field [23]. Finally, in 2015 Fernández, López, del Jesus, and Herrera revisited the topic of EFSs by presenting a complete taxonomy of the existing proposals, and also posing some new trends and challenges to suggest some potential research directions [29].

B. EFSs: Current taxonomy

As previously explained at the beginning of this paper, any EFS is developed on top of an FRBS. In this way, the components of the FRBS are learned or optimized using an evolutionary process commonly taken from available data, as illustrated in Figure 1. The final goal is being able to contextualize the behavior of these systems in a given scenario.

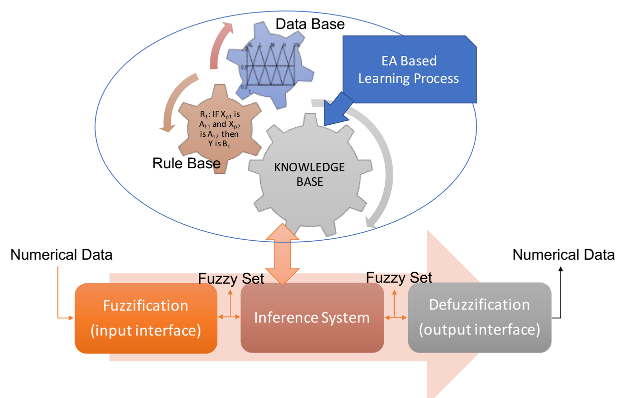

Fig. 1. How EFS are built on top of an FRBS. Inspired from [29]

The main reason for using EAs as a tool for the design of FRBSs is related to the possibility of addressing the optimization of their components as a search problem. Rule sets, membership functions, and many other features of an FRBS can be easily encoded inside a chromosome. The optimization procedure can be viewed from a double perspective: learning and/or tuning. In addition, the classical trade-off between accuracy and interpretability must be taken into account [30], [31]. For this task, the use of an MOEA is probably the option that is best suited. Finally, different ways of representing fuzzy sets may be considered as another aspect to be embedded in the optimization approach.

Taking all these aspects into account, in [29] authors proposed a complete taxonomy of EFSs, which is illustrated in Figure 2. Specifically, three large groups were highlighted depending on several aspects, such as how the FRBSs' elements are optimized, the trade-off among the different learning criteria, and the use of new fuzzy set representations.

In what follows, a brief description of the EFS approaches enumerated in the taxonomy is provided. First, the learning and tuning of FRBSs is introduced. Next, the use of multi-objective optimization in this framework is presented. Finally, several comments are given on the parameterized construction for new fuzzy representations.

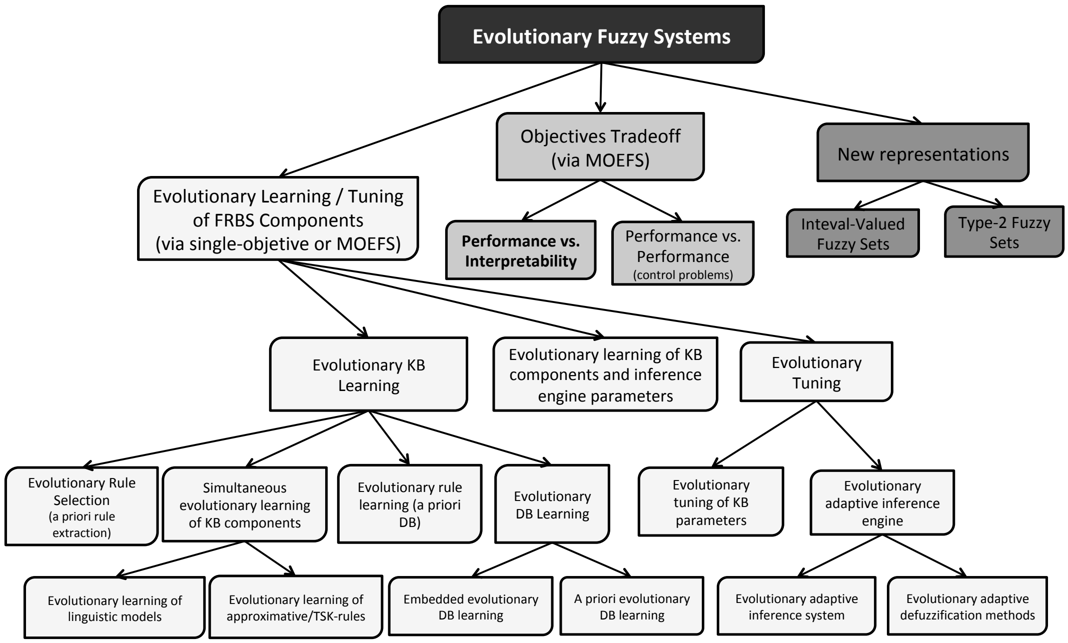

Fig. 2. Evolutionary Fuzzy Systems Taxonomy. First presented in Fernández et al. [29]

1) Evolutionary learning and tuning of FRBSs' components: When using an FRBS for modeling a given problem, researchers and practitioners must decide whether a simple system is enough for the existing requirements, or if a more complex computational solution is required. In such a case, the use of an EFS is mandatory in order to achieve a robust and accurate model.

There are two main alternatives for developing EFSs: (a) using EAs in the learning procedure of FRBSs and (b) using EAs for tuning the elements of FRBSs. Specifically, learning can be carried out either on the KB components or in conjunction with the inference engine parameters. The post-processing tuning is devoted to refining a preliminary definition of the FRBSs. In the following, a short description of the different approaches for both alternatives is provided.

a. *Evolutionary learning*. The specialized literature distinguishes between two main cases, depending on whether the learning is applied just to the KB or to the inference engine parameters as well.
   - *Evolutionary KB learning*. Four different learning options have been proposed to obtain the KB:
      i. *Evolutionary rule selection.* The goal is to remove useless rules in the final RB, i.e. those that may degrade the FRBS accuracy. Thus, a compact and accurate subset of fuzzy rules is intended to be the output for these methods.
      ii. *Simultaneous evolutionary learning of KB components.* As its name suggests, several FRBS elements may be obtained at once. The hitch in this case is that a larger search space is to be considered, implying a slower and harder learning process.
      iii. *Evolutionary rule learning.* This is by far the preferred approach in the specialized literature. Starting from a predefined DB (in most cases, composed of equally distributed fuzzy partitions), the fuzzy rules

of the RB are generated by means of the evolutionary process.

iv. *Evolutionary DB learning.* In this case, several parameters of the DB are considered. This includes the granularity degree, the shape of the membership functions, and the scaling functions, among other DB components. This DB learning can be carried out either after the RB is obtained or at the same time.

- *Evolutionary learning of KB components and inference engine parameters.*

  This comprises a special case, in which both the adaptive inference engine and the KB components are optimized. The main idea is to achieve the best synergy between the former elements, including both in a simultaneous learning process.

b. *Evolutionary tuning*. Tuning involves an *a-posteriori* optimization of the DB or the inference engine parameters. The RB is initially obtained by using a predefined DB and inference engine. Two different approaches have been proposed:

i. *Evolutionary tuning of KB parameters.* The parameters of the fuzzy sets are tuned in the evolutionary process.

ii. *Evolutionary adaptive inference engine*. It is divided into two groups depending on whether it is applied to the inference system or the defuzzification method. In the former case, the final objective is to obtain a higher cooperation among fuzzy rules by acting on the inference engine. The linguistic rule interpretability is maintained as the DB remains homogeneous. In the latter case, if a weighted average operator is used in the defuzzifier, its parameters can be optimized by means of EAs.

2) Objectives trade-off: Approaches for jointly optimizing several objectives: As seen in Section I, there is a growing interest in developing FRBSs that are both accurate and interpretable [30], [31]. However, this goal is not easy to achieve, as these criteria are usually in conflict. A smart solution is to carry out the learning procedure by means of MOEAs [22]. MOEAs generate a set of FRBSs with different trade-offs between the different learning objectives being considered in the optimization process. This solution is known as MOEFS [23], which can consider any metric of performance to carry out the optimization of the FRBSs, namely the cost, or the simplicity or comprehensibility, among others.

A short description of these techniques based on the multi-objective nature of the problem is given in what follows. In case of using MOEFSs for learning or tuning FRBS components, the reader should refer to those models introduced in the previous section.

- *Performance vs. Interpretability*. Fuzzy systems in general, and fuzzy linguistic models in particular, are well suited to understanding the nature of the problem that they represent. For this reason, it is quite important to consider different interpretability measures as an estimation of the FRBS potential.

  An FRBS is not interpretable *per se*. Instead, there are many different issues which must be taken into account in order to obtain a human interpretable structure. Among others, the rule base compactness or the semantic comprehensibility of the fuzzy partitions should be considered. Hence, the ML method used to learn the fuzzy model, classifier, or decision support systems must be properly designed to obtain the desired trade-off for the problem at hand.

Specifically, two different options can be taken into account. On the one hand, those which are based on the simplicity could be considered, i.e. the dimensionality, of the system (the simpler, the better). On the other hand, semantic-based metrics could be considered, i.e. the comprehensibility of the derived system.

- *Performance vs. Performance*. When addressing a control system problem, there are more constraints that must be taken into account than there are for standard data mining tasks. Specifically, there is a need to obtain an efficient controller, with proper stability and, if possible, to have a compact and interpretable structure. Therefore, the use of MOEAs for designing fuzzy controllers has become a very successful approach.

  In this case, both the structure of the controller and its parameters must be obtained. This implies a direct adaptation for all methods previously introduced for general EFSs. Specifically, a post-processing tuning step is the most common approach due to its simplicity and reduced search space.

3) New representations: fuzzy sets extensions: In several applications, where the degree of uncertainty is very high, traditional fuzzy sets are not able to provide an appropriate representation of the hidden knowledge. To cope with this aspect, some extensions such as type-2 fuzzy sets and interval-valued fuzzy sets are used instead.

Being special cases of fuzzy set representations, there is no established way of obtaining their parameter values and/or their structure. For this reason, EAs are a proper tool for the design strategy and/or for optimizing these fuzzy models. For example, the parameters of type-2 fuzzy sets may be optimized from a given standard fuzzy set, or by considering the direct generation from data. Recent approaches are aimed at tuning rules and conditions using this kind of fuzzy representation [32], [33].

III. What Evolutionary Fuzzy Systems Should Be Used For?

In the introduction of this manuscript, several interesting properties of fuzzy systems were highlighted. In short, they provide two interesting features making them very useful for knowledge representation tasks:

1) On the one hand, the inherent interpretability of the system. This refers to two different aspects. First, the intrinsic comprehensibility associated with the use of a simple description mechanism in the form of fuzzy linguistic rules, very close to natural language. And second, the comprehensibility and understandability of the rule-based system and the inference procedure.
2) On the other hand, the robustness of adapting to and learning from complex problems, i.e. to model scenarios which are difficult to represent with other types of paradigms. In particular, it is a very interesting tool to apply when users must deal with lack of data or uncertainty in the definition of the input data.

In regards to the aspects associated with the interpretability of FRBSs, first the “cointension” term must be highlighted, defined between fuzzy sets and regular concepts [34]. In particular, Lofti A. Zadeh emphasized that this semantic cointension was the key for understanding the success of the application of fuzzy models, namely the importance of the human component in the data science process [15]. This is what can be referred to as human-centric modeling and decision making, which implies the need to provide descriptive models and decision support systems able to comprehend the underlying human mental processes in order to manage the information in a more human-oriented style [35]. Human-centric models and descriptive support systems allow the designer to use the comprehensibility of the designed solutions to both understand the underlying human reasoning processes and uncover knowledge about the system at hand, as well as to enhance the problem solving [15].

Although the use of linguistic labels provides a solid basis for achieving both facets, many model induction techniques that use fuzzy systems can impose certain interpretability constraints (linguistic structure, rule length, and rule set size, among others) in the search for greater accuracy. It is at this point where EFS algorithms can be used to find a good trade-off between interpretability and accuracy [30], [31].

As previously discussed in detail in Section II-B, this can be done in one out of two phases. On the one hand, in the construction of the model itself, i.e. during the learning process. On the other hand, *a posteriori*, i.e. once the model has been obtained, a component tuning can be carried out. In both cases, the properties associated with EAs in regards to the coding of different types of information provide a clear benefit: *flexibility* [29]. In fact, they permit optimizing from simple parameters of fuzzy systems to complete sets of rules. The practitioner will therefore be able to directly pair the necessary interpretability constraints while generating a model with a high predictive power [30], [31].

Analyzing these properties in detail, the key ability of EFS is to obtain an ML system incorporating the fundamental property of what was defined as XAI [12]. It is not only about understanding the composition of the model, but also about the fact that the system is able to explain the user the process it followed to make the output decision. Therefore, the synergy between fuzzy systems and the ability of EAs to learn them makes EFSs a very appealing tool for a large number of problems. As an example of these developments, one may refer to the EFS design that allows the fusion mechanism for a classifier ensemble to be interpreted by the human designer thanks to a hierarchical fuzzy rule-based structure [36].

Focusing on the robustness properties associated with EFSs, the first aspect to be taken into account is their effectiveness in extracting information that resides in small datasets with low density in the domain, the so-called "lack of data" problem [37]. The reason is simple: by defining the universe of discourse of the fuzzy variables along the overall domain of the attributes that represent the problem, an integral coverage of the input space is allowed. Furthermore, when there is some overlap between the fuzzy sets, a smoother transition among the modeled information granules is expected to be obtained.

In addition to the lack of data, the benefits of EFSs in handling imprecise and uncertain data is also of high priority. In these cases, the flexibility of the definition of fuzzy partitions as well as the membership functions must be taken into account. In this sense, it is natural to use different extensions to the traditional type-1 fuzzy sets to add an extra level of freedom in the representation. However, defining the exact values according to the problem may not be a trivial task performed by a human operator. Thus, it is of great interest to be able to rely on the use of EAs to learn or adjust the components of the fuzzy DB.

It should be stressed that social network analysis or social mining are application areas that may benefit the most from EFSs [38]. It has become a hot topic in the last few years due to the increasing of social media interactions. Corporations and academia are very interested in conceptualizing, modeling, analyzing, explaining, and predicting these relationships. There is a natural connection between graph theory, on which social network analysis is based, and fuzzy set theory. This allows providing an easier and more robust way to express relationships among nodes in these networks. Furthermore, some research has already discussed the theoretical and conceptual models for social data based on fuzzy sets [39], [40].

Another area in which the inherent uncertainty of the data imposes a difficult restriction, and also requires interpretable models in order to be truly useful to the end-users, is finance. In this environment, the comprehension of how inputs and outputs are related to each other is crucial in order to be able to make operative and strategic decisions. Therefore, EFSs have been successfully applied to many financial domains [41].

Decisions in medical applications are also considered to be critical. Therefore, any action taken by experts must be taken confidently. This implies that any decision support system used in this context must be trustworthy and transparent. In other words, it must explain to both the doctor and the patient, the reasons behind a particular diagnosis. In this sense, an EFS based on fuzzy linguistic rules might be the proper choice [42].

Finally, current solutions focused on intrusion detection systems are high priority. The reason for this is clear, namely the vast use of information systems and the need of establishing security policies and rules that allow an undesired system access to be discriminated. In particular, EFS-based approaches are very interesting for several reasons. First of all, this type of problem has a common structure. Indeed, they are described by numeric data and therefore crisp thresholds can lead to low detection accuracy. Additionally, the boundary between legit and abnormal behavior is inherently fuzzy. In other words, small changes in an intrusion behavior may not be recognized, whereas a small deviation in a normal profile can generate a false alarm. In accordance to the former, several relevant EFS approaches based on fuzzy linguistic variables may be found in the specialized literature [43].

IV. Where are Evolutionary Fuzzy Systems Going? Future Prospects

From Section II, the reader might acknowledge that EFS-based methods have come a long way since the pioneering proposal of the first EFS-based methods more than 25 years ago. Extensive research has been carried out in this field, mainly due to the versatility associated with the learning and tuning of the different components of the FRBS.

Now, researchers and practitioners must look ahead to discern what the future objectives and challenges in the EFS field could be. The open directions for novel research are particularly associated with two basic pillars: (1) efforts focused on the novel optimization of the internal components of the FRBS (Section IV-A); and (2) analysis of the application of EFSs to emerging work scenarios in data science (Section IV-B). Finally, the particular characteristics of this big data scenario imply paying special attention to the design and development of EFSs (Section IV-C).

A. *Optimization of novel FRBS' components*

There is a wide variety of FRBS elements whose values can be optimized using an EFS. Enumerating some possibilities, we should stress the choice of relevant inputs, scaling factors, membership functions, shape functions, granularity, fuzzy rules, inference parameters, number of rules, and so on. In addition, as previously described the power of EAs permits the joint learning of several of these components so as to obtain a much more robust FRBS.

However, it must be highlighted once again that for current applications it is not useful to generate models with excessively complex components. In other words, special attention must be paid to maintain the comprehensibility of the models. For this reason, the evolution of the multiple components of an FRBS must always be carried out considering the semantic properties and simplicity of the obtained system.

B. Emerging data science scenarios for EFS

Data science is a quite recent field of study, and it is still rapidly expanding. New problems and practical applications arise and require the development of robust techniques to address new complex paradigms. Specifically, novel problems are usually characterized by: i) uncertainty, the available information is often imprecise, uncertain, noisy, or there might be an acute lack of information; the objective of the modeling or decision problem is ambiguous and the problem structure might be loosely specified; and the problem environment may be not stationary but changing; ii) the impressively increasing problem dimensions, requiring prohibitive computational times to achieve problem tractability; and iii) the need to provide descriptive models and decision support systems able to comprehend the underlying human mental processes to manage the information in a more human-oriented style (human-centric modeling and decision making) [44]. EFSs have excelled in many different scenarios and, in accordance with this general good behavior, a bright future for their use in most of the incoming data science areas can be predicted.

A first case-study is related to non-standard classification problems, such as those based on multi-label and multi-instance learning. In the former case, the model must classify a query instance in a set of non-exclusive categories [45]. In the latter, incomplete information about the classes or instances implies a handicap during the categorization [46]. Thus, instances are coupled into a single "bag", being unaware of which is the one that provides the true label. Finally, both problems can be combined adding a higher degree of complexity to the design of future solutions. Taking into account this type of relationship between labels and/or instances, the nature of both problems is well suited for the use of fuzzy sets and systems, and thus for EFSs. However, at present just a few attempts have been made to solve this task using these tools [47].

The novel topic related to ordinal and monotonic classification [48] should be noted. This case study comprises those problems where both the input attributes and/or the class have a monotonicity constraint. The first study that extracts fuzzy rules satisfying these constraints without the need of a preprocessing stage can be found in [49], where authors incorporate several mechanisms based on monotonicity with EFS algorithms.

Supervised descriptive rule discovery [50], and in particular subgroup discovery [51], [52], is also another interesting area of study. The goal is to locate subgroups which are statistically "most interesting" for the user. The obtained model must fulfill some properties such as simplicity (understandable structure with few variables), and both a high significance and support. Another particular case of this scenario is the so-called Emerging Pattern Extraction [53], in which frequency changes significantly from one dataset to another. In this context, the use of MOEFSs has shown that obtained rules allow the descriptions of the emerging phenomena to be simpler than those in the state of the art [54].

A topic that has gained much attention in the research community of data science is multi-view learning [55], where examples are described by multiple distinct feature sets. This is related to the way several current-day problems are defined, such as multimedia-content, web page classification, and bioinformatics. The approaches which address multi-view learning via data integration are quite diverse.

Another interesting area of study is semi-supervised learning. It is based on problems in which only a subset of the instances are labeled and the algorithm can modify the output of the training data [56]. Considered as an extension to unsupervised learning, clustering techniques have received much attention. In this context, the use of an FRBS can represent a sort of linguistic description of the dissimilarity relation among patterns [57], thus helping the recognition of the clusters with fewer data.

Finally, there are novel domains of application arising everyday that share problems similar to the ones already solved in other domains. Therefore, it can be of interest to exploit previously acquired knowledge to manage new tasks in a quick and effective way. This is the premise of the area of research known as transfer learning [58]. Taking advantage of fuzzy learning methods for this task is evident, regarding the uncertainty that is found in these dynamic environments [59]. They also allow leveraging knowledge from some referring scenes when there is little data available [60].

These are some of the new data science scenarios without the aim of being very exhaustive.The previous analisys ilustrates that there are not enough scientific studies using EFSs for such novel and significant topics. Therefore, this must be regarded as a “call-to-action” for current researchers in the area of EFSs to open the way in such promising lines of study, from both the theoretical and practical points of view.

C. EFSs in big data: a significant topic for the near future

One of the hottest topics for current research is related to data science and big data problems [4]. An in-depth analysis of the current state of this framework was carried out in both [61] and [62], where authors reported the good properties of fuzzy systems when dealing with such types of applications. However, focusing on the case of EFSs for big data, so far few studies have been developed in this area of research [63]–[65].

The main reason for the lack of proposals in the specialized literature is the difficulty in reaching solutions that are scalable within the EFS paradigm. Indeed, this constraint is associated with the computation of an accuracy-based fitness function from data at the core of the evolutionary procedure. In other words, it is not straightforward to develop methodologies that allow the whole dataset to be handled in reasonable time.

For this reason, the development of learning and tuning methods must be redirected within a distributed environment. To this end, MapReduce has established as a *de-facto* solution to simplify the implementation of these techniques [66].

It is basically an execution environment which lays over a distributed and fault tolerant file system, HDFS being the most common option. By means of two simple functions, Map and Reduce, any implementation can be automatically parallelized in a transparent way for the programmer. The first function allows operating on independent “chunks” of data, by applying the same procedure. The second merges the outputs of the Map functions to produce a single final output.

According to the new MapReduce paradigm of distributed programming, any ML solution can be categorized into two main types [67]:

1) Local approaches (also known as approximate models) that work directly on the distributed chunks of data by creating partial models that are then aggregated.
2) Global approaches (also known as exact models) that iterate over all examples to generate the model or build the system iteratively by optimal merging.

By analyzing the properties of each of these types of methodologies some very interesting insights could be achieved.

Local approaches are *a priori* easier to develop than the original algorithm would be embedded in the different Map tasks. However, it would imply a greater effort in the merging phase of the independent sub-models (Reduce function). Another advantage is an expected efficiency increase when augmenting the number of partitions. The main hitch in this case would be to work with a smaller number of data in each of the sub-processes. Fortunately, the good behavior of FRBSs and EFSs with respect to such an eventuality has been previously highlighted [62] (see Section III).

Global approaches, on the other hand, have as their main virtue the learning of more robust models. This is a theoretical issue as they comprise the analysis of the complete set of data. However, this is precisely their major disadvantage. Specifically, more effort must be devoted to design the methodology to meet the conditions of correctness. In addition to the former, the efficiency can also be reduced as the process must be iterated several times.

A very interesting option in this case is to take advantage of the features of novel environments such as Spark [68]. It is a distributed computing platform that provides a memory-intensive scheme, being very suitable for ML algorithms [67]. Specifically, it supports very versatile data structures, together with wide range of operations for transforming them, such as filtering, grouping, or set operations, among others.

When referring to big data constraints for the development of novel ML approaches, the efforts are mainly focused on finding scalable solutions considering only one-side of the *Volume*, namely the "Big Instance Size". However, there is another significant issue in this regard, which is known as the "Big Dimensionality" explosion [69]. There are diverse ways to face this problem, including smart feature selection [70], fuzzy ensemble models based on bagging for vertically partitioning the training set [71], [72], and the use of fuzzy decision trees that internally consider a feature ranking mechanism [73].

As it has been discussed throughout this section, the task of addressing big data problems with EFSs is still far from being fully covered. Many state-of-the-art learning and tuning approaches need to be redesigned and translated to the new MapReduce paradigm. The objective is to take advantage of the power and quality that these algorithms have already shown, but using the new applications and current-day data sets. Furthermore, researchers must be aware of this circumstance to be more ambitious and develop new robust methodologies by taking into account the new functionalities available in the current programming environments. As a final remark, the difficulties associated with this framework imply the development of much more ingenious and effective ideas to achieve greater milestones in this area of research.

V. Remarks On the Need for Interpretable and Explainable Artificial Intelligence

Throughout this paper, it has been stressed that there are two main criteria any practitioner must consider when determining what kind of ML techniques should be used to address a specific regression, classification, or decision making problem: accuracy and interpretability/explainability. Of course, the ideal situation would be to obtain jointly high degrees of both but, since they are in conflict, this is a complex task. Figure 3 shows this fact with a trade-off for some well-known ML techniques and the current challenge to increase the explanaibility without reducing accuracy. For the most interpretable models the challenge is to increase both axis.

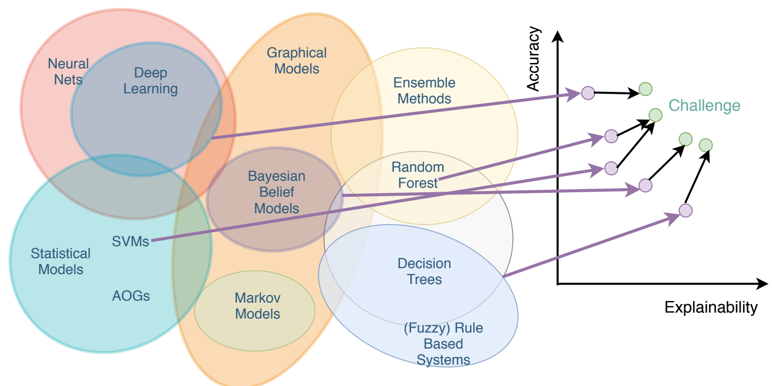

Fig. 3. Accuracy vs. Explainability: Venn diagram for several ML models (partially inspired from [74]).

In practice, one of the two properties usually prevails over the other (use of black-box vs. white- or gray-box models/classifiers). Driven by the need for accuracy, we are witnessing an emerging trend to embrace "black-box" solutions, in particular those based on neural networks and mainly Deep Learning (DL) [75], [76]. In short, these systems are based on a number of layers of fully-connected neurons. The first layers are devoted to extracting simpler attributes from the data, which are then combined in the subsequent layers to form more complex and thus more representative attributes. In spite of the impressive capacity of these black-box solutions to obtain accurate models, this virtue is often associated with a higher system complexity. Designers are often unable to understand how the systems work, and also to decipher why they produce a certain output. In fact, the effectiveness of the existing ML methods could be limited by the machines' inability to explain its thoughts and actions to human users in some application fields, and thus may lead to unsafe and incorrect decisions[1].

Interpreting and explaining deep networks is a young and emerging field of research. In [17] can be found a study of techniques for interpreting complex machine learning models, with focus on DL. Most of the proposals have focused on post-hoc interpretability but greater efforts are required in the near future to understand the ML models.

In this scenario, rule based systems [77] allow to audit the extracted knowledge with a double objective. On the one hand, obtaining a clear explanation of the cognition process carried out by the system. On the other hand, being able to trust in the description of the rules and their relationship with the problem that is aimed to be solved. In particular, EFSs combine the ability to represent knowledge in natural way for human understanding (with fuzzy rules), the strength of fuzzy reasoning and the ability of EAs for search in complex and ill-defined problems. However, it must be emphasized that FRBSs must remain simple and understandable, since they are not interpretable

[1] https://www.statnews.com/2018/07/25/ibm-watson-recommended-unsafe-incorrect-treatments/

per se [78]. It is important to take account of different issues in order to obtain FRBSs that represent knowledge easily understood by humans. Among others, the rule base compactness or the semantic comprehensibility of the fuzzy partitions must be stressed. Moreover, the EFSs must be properly designed to obtain the desired trade-off between accuracy and explainability for the problem at hand. When in the EFS design more attention is paid to the accuracy than to explainability, the skills of the fuzzy system obtained are hardly comparable with other preferable and more complex solutions such as DL.

DL is a powerful paradigm that can effectively capture relevant features, in particular for those problems from which a higher level of abstraction is needed to describe the hidden knowledge [76]. In any way, DL solutions must not be considered to be rivals for FRBSs and EFSs, but rather as complementary approaches, each one with their strengths and drawbacks. A promising step forward the achievement of this objective could be to make use of the good interpretability properties of EFSs in conjunction with DL.

There are some interesting examples of the positive synergy between both paradigms. First, a fuzzy classifier based on stacking is proposed in [79]. It embeds a zero-order TSK FRBS with a limited number of linguistic labels within a neuron-layered representation. Thus, each layer becomes a single model within the "ensemble". After the learning stage using a DL procedure, rules can be simply extracted from each "neuron" to keep the original interpretability. A different approach is proposed in [80]. It comprises a hierarchical system composed of a fuzzy rule layer, to manage the ambiguity of the data, and a DL layer, to reduce the noise and to create higher order variables for a more accurate representation. Then, both parts are fused leading to a more robust classification. Finally, an interpretable structure for DL networks was also proposed in [81]. To achieve the required explainability, zero-order TSK fuzzy linguistic rules are encoded and learned in the final layer of the net to perform the final classification, just after the feature extraction.

In summary, EFSs are a very significant tool in many fields of application where the explainability of the decision must be taken into account. The reader should refer to areas such us medical diagnosis, financial problems, or security systems, among others. It is straightforward to acknowledge that, in these scenarios, the same importance must be given to the accuracy/confidence of the output and the explainability of the decision made. In other words, ML models based on EFSs are a desirable choice when the output is intended to be trusted by the human user.

VI. Concluding Remarks

The world of data science has changed the way applications are approached. At present, the core of the model not only aims to achieve the highest possible accuracy, but also to make it explainable for researchers and practitioners. In this sense, ML methods based on EFSs preserve the original essence of comprehensibility exposed by Zadeh, also boosting their modeling abilities. It is straightforward to acknowledge that this provides several advantages over other paradigms toward handing XAI learning models including transparency, understanding and comprehensibility.

Throughout this work, a variety of perspectives for understanding the virtues of EFSs have been identified. A series of "4 W questions" (*why, when, what for, and where to*) have been posed. The objective was that the answers provide some insight into the capabilities that EFSs have shown when being adapted to different research areas, and to promote new developments in the discipline.

It has enable us to notice the lack of unanimity in the literature about XAI concepts, their formal definition, quantitative measures and relations among them. There are currently numerous coexisting approaches to interpretability / explainability. For this reason there is a need for a debate within the AI, ML and fuzzy communities to standarise and to assess these concepts and the auxiliary properties than let to interpret and explain properly the AI models and outputs.

ACKNOWLEDGMENTS

The authors would like to ennoble Prof. Lofti A. Zadeh's figure, both from a personal and a professional viewpoint. They would like to acknowledge the pioneering development of the fuzzy sets and systems research area and the strong positive influence Prof. Zadeh played for its development during more than 50 years. The authors are very proud of having met and collaborated with Prof. Zadeh in many different activities both at the University of Granada and the European Centre for Soft Computing. They had the chance to recognize how Lotfi was not only an impressive researcher with a visionary mind but, and even more important, an outstanding human being. He will be missed.

This work have been partially supported by the Spanish Ministry of Science and Technology under projects TIN2015-68454-R, TIN2015-67661-P, and TIN2017-89517-P, including European Regional Development Funds.

REFERENCES

[1] F. Bacon, *Meditationes sacrae*. Excusum impensis Humfredi Hooper, 1597.

[2] M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects," *Science*, vol. 349, no. 6245, pp. 255–260, 2015.

[3] V. Dhar, "Data science and prediction," *Communications of the ACM*, vol. 56, pp. 64–73, 2013.

[4] A. Fernández, S. Río, V. López, A. Bawakid, M. J. del Jesus, J. Benítez, and F. Herrera, "Big Data with cloud computing: An insight on the computing environment, MapReduce and programming framework," *WIREs Data Mining and Knowledge Discovery*, vol. 4, no. 5, pp. 380–409, 2014.

[5] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, pp. 338–353, 1965.

[6] H. Ishibuchi, T. Nakashima, and M. Nii, *Classification and Modeling with Linguistic Information Granules*. Springer, 2005.

[7] E. H. Mamdani and S. Assilian, "An experiment in linguistic synthesis with a fuzzy logic controller," *International Journal of Man-Machine Studies*, vol. 7, no. 1, pp. 1–13, 1975.

[8] A. E. Eiben and J. E. Smith, *Introduction to evolutionary computation*. Berlin, Germany: Springer-Verlag, 2003.

[9] J. H. Holland, *Adaptation in Natural and Artificial Systems*. Ann Arbor, MI, USA: University of Michigan Press, 1975.

[10] A. Freitas, *Data Mining and Knowledge Discovery with Evolutionary Algorithms*, ser. Natural Computing Series. Berlin Heidelberg: Springer, 2002.

[11] W. Samek, T. Wiegand, and K. R. Müller, "Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models." *CoRR*, vol. abs/1708.08296, 2017. [Online]. Available: http://www.arxiv.org/abs/1708.08296

[12] D. Castelvecchi, "Can we open the black box of AI?" *Nature*, vol. 538, no. 7623, pp. 20–23, 2016.

[13] S. F. Bryce Goodman, "European union regulations on algorithmic decision-making and a "right to explanation"," *AI Magazine, Vol 38, No 3*, vol. abs/1606.08813v3, 2017. [Online]. Available: https://arxiv.org/abs/1606.08813v3

[14] F. Doshi-Velez and B. Kim, "Towards a rigorous science of interpretable machine learning," *CoRR*, vol. abs/1702.08608, 2017. [Online]. Available: https://arxiv.org/abs/1702.08608

[15] M. Gleicher, "A framework for considering comprehensibility in modeling," *Big Data*, vol. 4, no. 2, pp. 75–88, 2016.

[16] Z. C. Lipton, "The mythos of model interpretability," *CoRR*, vol. abs/1606.03490, 2017. [Online]. Available: https://arxiv.org/abs/1606.03490

[17] G. Montavon, W. Samek, and K. R. Müller, "Methods for interpreting and understanding deep neural networks," *Digital Signal Processing*, vol. 73, pp. 1–15, 2018.

[18] M. W. Craven, *Extracting comprehensible models from trained neural networks*. Thesis. University of Wisconsin, 1996.

[19] R. S. Michalski, "A theory and methodology of inductive learning," *Artificial Intelligence*, vol. 20, pp. 111–161, 1983.

[20] M. Ribeiro, S. Singh, and C. Guestrin, ""Why should I trust you?" explaining the predictions of any classifier," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 1135–1144.

[21] O. Cordon, F. Herrera, F. Hoffmann, and L. Magdalena, *Genetic fuzzy systems. Evolutionary tuning and learning of fuzzy knowledge bases*. Singapore, Republic of Singapore: World Scientific, 2001.

[22] C. A. Coello-Coello, G. Lamont, and D. van Veldhuizen, *Evolutionary algorithms for solving multi-objective problems*, 2nd ed., ser. Genetic and Evolutionary Computation. Berlin, Heidelberg: Springer, 2007.

[23] M. Fazzolari, R. Alcala, Y. Nojima, H. Ishibuchi, and F. Herrera, "A review of the application of multi-objective evolutionary systems: Current status and further directions," *IEEE Trans. Fuzzy Syst.*, vol. 21, no. 1, pp. 45–65, 2013.

[24] C. Karr, "Genetic algorithms for fuzzy controllers," *AI Expert*, vol. 6, no. 2, pp. 26–33, 1991.

[25] M. Valenzuela-Rendon, "The fuzzy classifier system: A classifier system for continuously varying variables," in *4th International Conference on Genetic Algorithms (ICGA'91)*, 1991, pp. 346–353.

[26] P. Thrift, "Fuzzy logic synthesis with genetic algorithms," in *Proceedings of the 4th International Conference on Genetic Algorithms (ICGA'91)*, 1991, pp. 509–513.

[27] D. Pham and D. Karaboga, "Optimum design of fuzzy logic controllers using genetic algorithms," *Journal of Systems Engineering*, vol. 1, pp. 114–118, 1991.

[28] O. Cordon, "A historical review of evolutionary learning methods for Mamdani-type fuzzy rule-based systems: Designing interpretable genetic fuzzy systems," *Int. J. Approx. Reason.*, vol. 52, no. 6, pp. 894–913, 2011.

[29] A. Fernandez, V. Lopez, M. J. del Jesus, and F. Herrera, "Revisiting evolutionary fuzzy systems: Taxonomy, applications, new trends and challenges," *Knowlegde Based Systems*, vol. 80, pp. 109–121, 2015.

[30] J. Alonso, L. Magdalena, and G. González-Rodríguez, "Looking for a good fuzzy system interpretability index: An experimental approach," *International Journal of Approximate Reasoning*, vol. 51, pp. 115–134, 2009.

[31] M. J. Gacto, R. Alcala, and F. Herrera, "Interpretability of linguistic fuzzy rule-based systems: An overview of interpretability measures," *Inform. Sciences*, vol. 181, no. 20, pp. 4340–4360, 2011.

[32] J. Sanz, D. Bernardo, F. Herrera, H. Bustince, and H. Hagras, "A compact evolutionary interval-valued fuzzy rule-based classification system for the modeling and prediction of real-world financial applications with imbalanced data," *IEEE Trans. Fuzzy Syst.*, vol. 23, no. 4, pp. 973–990, 2015.

[33] K. Tai, A.-R. El-Sayed, M. Biglarbegian, C. I. Gonzalez, O. Castillo, and S. Mahmud, "Review of recent type-2 fuzzy controller applications," *Algorithms*, vol. 9, no. 2, 2016.

[34] L. A. Zadeh, "Is there a need for fuzzy logic?" *Inform. Sciences*, vol. 178, no. 13, pp. 2751–2779, 2008.

[35] P. Guo and W. Pedrycz, Eds., *Human-Centric Decision-Making Models for Social Sciences*, ser. Studies in Computational Intelligence. Springer, 2014, vol. 502.

[36] K. Trawinski, O. Cordon, L. Sanchez, and A. Quirin, "A genetic fuzzy linguistic combination method for fuzzy rule-based multi-classifiers," *IEEE Trans. Fuzzy Syst.*, vol. 21, no. 5, pp. 950–965, 2013.

[37] V. Lopez, A. Fernandez, S. Garcia, V. Palade, and F. Herrera, "An insight into classification with imbalanced data: Empirical results and current trends on using data intrinsic characteristics," *Inform. Sciences*, vol. 250, no. 20, pp. 113–141, 2013.

[38] G. Bello-Orgaz, J. J. Jung, and D. Camacho, "Social big data: Recent achievements and new challenges," *Information Fusion*, vol. 28, pp. 45 – 59, 2016.

[39] S. Bastani, A. K. Jafarabad, and M. H. F. Zarandi, "Fuzzy models for link prediction in social networks." *Int. J. Intell. Syst.*, vol. 28, no. 8, pp. 768–786, 2013.

[40] M. Dragoni and G. Petrucci, "A fuzzy-based strategy for multi-domain sentiment analysis," *Int. J. Approx. Reason.*, vol. 93, pp. 59–73, 2018.

[41] M. Antonelli, D. Bernardo, H. Hagras, and F. Marcelloni, "Multiobjective evolutionary optimization of type-2 fuzzy rule-based systems for financial data classification," *IEEE Trans. Fuzzy Syst.*, vol. 25, no. 2, pp. 249–264, 2017.

[42] S. A. Mokeddem, "A fuzzy classification model for myocardial infarction risk assessment," *Applied Intelligence*, vol. 48, no. 5, pp. 1233–1250, 2018.

[43] S. Elhag, A. Fernández, A. Altalhi, S. Alshomrani, and F. Herrera, "A multi-objective evolutionary fuzzy system to obtain a broad and accurate set of solutions in intrusion detection systems," *Soft Comp.*, vol. in press, 2018.

[44] M. Chen, F. Herrera, and K. Hwang, "Cognitive computing: Architecture, technologies and intelligent applications," *IEEE Access*, vol. 6, pp. 19 774–19 783, 2018.

[45] F. Herrera, F. Charte, A. J. Rivera, and M. J. del Jesús, *Multilabel Classification - Problem Analysis, Metrics and Techniques*. Springer, 2016.

[46] F. Herrera, S. Ventura, R. Bello, C. Cornelis, A. Zafra, D. S. Tarragó, and S. Vluymans, *Multiple Instance Learning - Foundations and Algorithms*. Springer, 2016.

[47] S. Vluymans, D. Sanchez-Tarrago, Y. Saeys, C. Cornelis, and F. Herrera, "Fuzzy multi-instance classifiers," *IEEE Trans. Fuzzy Syst.*, vol. 24, no. 6, pp. 1395–1409, 2016.

[48] W. Kotlowski and R. Slowinski, "On nonparametric ordinal classification with monotonicity constraints," *IEEE Trans. Knowl. Data Eng.*, vol. 25, no. 11, pp. 2576–2589, 2013.

[49] J. Alcala-Fdez, R. Alcala, S. Gonzalez, Y. Nojima, and S. Garcia, "Evolutionary fuzzy rule-based methods for monotonic classification," *IEEE Trans. Fuzzy Syst.*, vol. 25, no. 6, pp. 1376–1390, 2017.

[50] P. Novak, N. Lavrac, and G. Webb, "Supervised descriptive rule discovery: A unifying survey of contrast set, emerging pattern and subgroup mining," *J. Mach. Learn. Res.*, vol. 10, pp. 377–403, 2009.

[51] W. Klösgen, "Explora: A multipattern and multistrategy discovery assistant," in *Advances in Knowledge Discovery and Data Mining*. AAAI, 1996, pp. 249–271.

[52] F. Herrera, C. J. Carmona, P. González, and M. J. Del Jesus, "An overview on subgroup discovery: Foundations and applications," *Knowl. Inf. Syst.*, vol. 29, no. 3, pp. 495–525, 2011.

[53] G. Dong and J. Li, "Mining border descriptions of emerging patterns from dataset pairs." *Knowl. Inf. Syst.*, vol. 8, no. 2, pp. 178–202, 2005.

[54] A. Garcia Vico, C. Carmona, P. Gonzalez, and M. Del Jesus, "MOEA-EFEP: multi-objective evolutionary algorithm for the extraction of fuzzy emerging patterns," *IEEE Trans. Fuzzy Syst.*, vol. in press, 2018.

[55] J. Zhao, X. Xie, X. Xu, and S. Sun, "Multi-view learning overview," *Inf. Fusion*, vol. 38, no. C, pp. 43–54, Nov. 2017. [Online]. Available: https://doi.org/10.1016/j.inffus.2017.02.007

[56] I. Triguero, S. García, and F. Herrera, "Self-labeled techniques for semi-supervised learning: taxonomy, software and empirical study," *Knowl. Inf. Syst.*, vol. 42, no. 2, pp. 245–284, 2013.

[57] M. G. C. A. Cimino, B. Lazzerini, and F. Marcelloni, "A novel approach to fuzzy clustering based on a dissimilarity relation extracted from data using a TS system." *Pattern Recogn.*, vol. 39, no. 11, pp. 2077–2091, 2006.

[58] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, 2010.

[59] J. Shell and S. Coupland, "Fuzzy transfer learning: Methodology and application," *Inform. Sciences*, vol. 293, pp. 59–79, 2015.

[60] H. Zuo, G. Zhang, W. Pedrycz, V. Behbood, and J. Lu, "Fuzzy regression transfer learning in Takagi-Sugeno fuzzy models," *IEEE Trans. Fuzzy Syst.*, vol. 25, no. 6, pp. 1795–1807, 2017.

[61] A. Fernandez, C. Carmona, M. del Jesus, and F. Herrera, "A view on fuzzy systems for big data: Progress and opportunities," *International Journal of Computational Intelligence Systems*, vol. 9, no. 1, pp. 69–80, 2016.

[62] A. Fernandez, A. Altalhi, S. Alshomrani, and F. Herrera, "Why linguistic fuzzy rule based classification systems perform well in big data applications?" *International Journal of Computational Intelligence Systems*, vol. 10, pp. 1211–1225, 2017.

[63] I. Rodríguez-Fdez, M. Mucientes, and A. Bugarín, "S-FRULER: Scalable fuzzy rule learning through evolution for regression," *Knowl.-Based Syst.*, vol. 110, pp. 255–266, 2016.

[64] A. Ferranti, F. Marcelloni, A. Segatori, M. Antonelli, and P. Ducange, "A distributed approach to multi-objective evolutionary generation of fuzzy rule-based classifiers from big data," *Inform. Sciences*, vol. 415-416, pp. 319–340, 2017.

[65] F. Pulgar-Rubio, A. Rivera-Rivas, M. Pérez-Godoy, P. González, C. Carmona, and M. del Jesus, "MEFASD-BD: Multi-objective evolutionary fuzzy algorithm for subgroup discovery in big data environments - a mapreduce solution," *Knowl.-Based Syst.*, vol. 117, pp. 70–78, 2017.

[66] K. H. Lee, Y. J. Lee, H. Choi, Y. D. Chung, and B. Moon, "Parallel data processing with mapreduce: A survey," *SIGMOD Record*, vol. 40, no. 4, pp. 11–20, 2011.

[67] S. Ramírez-Gallego, A. Fernández, S. García, M. Chen, and F. Herrera, "Big data: Tutorial and guidelines on information and process fusion for analytics algorithms with mapreduce." *Information Fusion*, vol. 42, pp. 51–61, 2018.

[68] M. Hamstra, H. Karau, M. Zaharia, A. Konwinski, and P. Wendell, *Learning Spark: Lightning-Fast Big Data Analytics*. O'Reilly Media, Incorporated, 2015.

[69] Y. Zhai, Y. S. Ong, and I. W. Tsang, "The emerging "big dimensionality"," *IEEE Comput. Intell. Mag.*, vol. 9, no. 3, pp. 14–26, 2014.

[70] S. Ramírez-Gallego, H. Mouriño-Talín, D. Martínez-Rego, V. Bolón-Canedo, J. M. Benítez, A. Alonso-Betanzos, and F. Herrera, "An information theory-based feature selection framework for big data under apache spark," *IEEE Trans. Systems, Man, and Cybernetics: Systems*, vol. in press, pp. 1–13, 2018.

[71] P. Bonissone, J. Cadenas, M. Carmen Garrido, and R. Andrés Díaz-Valladares, "A fuzzy random forest," *Int. J. Approx. Reason.*, vol. 51, no. 7, pp. 729–747, 2010.

[72] K. Trawiski, O. Cordón, and A. Quirin, "On designing fuzzy rule-based multiclassification systems by combining furia with bagging and feature selection," *International Journal of Uncertainty, Fuzziness and Knowlege-Based Systems*, vol. 19, no. 4, pp. 589–603, 2011.

[73] A. Segatori, F. Marcelloni, and W. Pedrycz, "On distributed fuzzy decision trees for big data," *IEEE Trans. Fuzzy Syst.*, vol. 26, no. 1, pp. 174–192, 2018.

[74] D. Gunning, "Explainable artificial intelligence (xai)," Defense Advanced Research Projects Agency, DARPA/I20, 2017.

[75] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[76] Z. C. P. L. Q. Zhang, L.T. Yang, "A survey on deep learning for big data," *Information Fusion*, vol. 42, pp. 146–157, 2018.

[77] J. Fürnkranz, D. Gamberger, and N. Lavrac, *Foundations of Rule Learning*. Springer, 2012.

[78] L. I. Kuncheva, "How good are fuzzy if-then classifiers?" *IEEE Trans. Systems, Man, and Cybernetics, Part B*, vol. 30, no. 4, pp. 501–509, 2000.

[79] T. Zhou, F.-L. Chung, and S. Wang, "Deep TSK fuzzy classifier with stacked generalization and triplely concise interpretability guarantee for large data." *IEEE Trans. Fuzzy Syst.*, vol. 25, no. 5, pp. 1207–1221, 2017.

[80] Y. Deng, Z. Ren, Y. Kong, F. Bao, and Q. Dai, "A hierarchical fused fuzzy deep neural network for data classification." *IEEE Trans. Fuzzy Syst.*, vol. 25, no. 4, pp. 1006–1012, 2017.

[81] P. Angelov and X. Gu, "A cascade of deep learning fuzzy rule-based image classifier and SVM," in *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2017, pp. 746–751.