Knowledge-Based Systems 80 (2015) 109–121

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Revisiting Evolutionary Fuzzy Systems: Taxonomy, applications, new trends and challenges

Alberto Fernández [a,*], Victoria López [b], María José del Jesus [a], Francisco Herrera [b,c]

[a] Department of Computer Science, University of Jaén, Jaén, Spain
[b] Department of Computer Science and Artificial Intelligence, CITIC-UGR (Research Center on Information and Communications Technology), University of Granada, Granada, Spain
[c] Faculty of Computing and Information Technology – North Jeddah, King Abdulaziz University (KAU), Jeddah, Saudi Arabia

ARTICLE INFO

Article history:
Received 31 October 2014
Received in revised form 4 January 2015
Accepted 23 January 2015
Available online 4 February 2015

Keywords:
Evolutionary Fuzzy Systems
Multi-Objective Evolutionary Fuzzy Systems
Fuzzy Rule Based Systems
Taxonomy
New trends
Data Mining
Scalability
Big Data

ABSTRACT

Evolutionary Fuzzy Systems are a successful hybridization between fuzzy systems and Evolutionary Algorithms. They integrate both the management of imprecision/uncertainty and inherent interpretability of Fuzzy Rule Based Systems, with the learning and adaptation capabilities of evolutionary optimization. Over the years, many different approaches in Evolutionary Fuzzy Systems have been developed for improving the behavior of fuzzy systems, either acting on the Fuzzy Rule Base Systems' elements, or by defining new approaches for the evolutionary components.

All these efforts have enabled Evolutionary Fuzzy Systems to be successfully applied in several areas of Data Mining and engineering. In accordance with the former, a wide number of applications have been also taken advantage of these types of systems. However, with the new advances in computation, novel problems and challenges are raised every day. All these issues motivate researchers to make an effort in releasing new ways of addressing them with Evolutionary Fuzzy Systems.

In this paper, we will review the progression of Evolutionary Fuzzy Systems by analyzing their taxonomy and components. We will also stress those problems and applications already tackled by this type of approach. We will present a discussion on the most recent and difficult Data Mining tasks to be addressed, and which are the latest trends in the development of Evolutionary Fuzzy Systems.

1. Introduction

Fuzzy systems have become a very popular tool in the Computational Intelligence and Soft Computing area for solving different problems in Data Mining and engineering [123]. This is due to their good properties for representing uncertain knowledge and therefore for adapting more smoothly to the context in which they are working at. Usually, model structure in the form of Fuzzy Rule Based Systems (FRBSs) is considered. FRBSs constitute an extension to classical rule-based systems, because they deal with "IF-THEN" rules, but its antecedents and consequents are composed of fuzzy logic statements, instead of classical ones. Additionally, in the case of fuzzy sets with linguistic labels, the output system has a higher interpretability degree for the expert to understand the working procedure of the former [73], and the inner details of the problem characteristics [62].

At the beginning of the 1990s, the combination between fuzzy systems and Evolutionary Computation was studied. The main idea behind this synergy was to take advantage of the optimization capabilities of Evolutionary Algorithms (EAs) [50] for improving the accuracy of fuzzy systems. Specifically, this was carried out by either by performing an automatic definition of the FRBSs, or by tuning some of the elements of their structure, i.e. Data Base (DB), Rule Base (RB), or inference system.

This hybridization had led to Evolutionary Fuzzy Systems (EFSs), which comprises an extension of the traditional and well-known Genetic Fuzzy Systems [39,75]. The former term was due to the use of Genetic Algorithms (GAs) as the primary part of this synergy. In this paper we focus on a generic type of evolutionary techniques. Indeed, EFSs have been recently extended by using Multi-Objective Evolutionary Algorithms (MOEAs) [32], considering multiple conflicting objectives, instead of a single one. These specific types of approaches are known as Multi-Objective Evolutionary Fuzzy Systems (MOEFSs), and they have become an important part of the more general EFSs [53].

In this paper we will revisit the different EFSs approaches that have been proposed in the specialized literature. Furthermore,

* Corresponding author. Tel.: +34 953 213016; fax: +34 953 212472.
E-mail addresses: alberto.fernandez@ujaen.es (A. Fernández), vlopez@decsai.ugr.es (V. López), mjjesus@ujaen.es (M.J. del Jesus), herrera@decsai.ugr.es (F. Herrera).

http://dx.doi.org/10.1016/j.knosys.2015.01.013
0950-7051/

we will analyze the latest trends and challenges for the development of new EFS methods. With this aim, we will divide this work into five main parts as follows:

- We will first introduce a complete taxonomy regarding three different aspects, i.e. the learning and/or optimization of the FRBSs' elements, the multi-objective approaches for achieving the tradeoff among different objectives, and the development of tuning algorithms for the definition of novel fuzzy representations. This will allow us to have a global view of the organization of the EFS models, so that we can have a better understanding of the evolution and characteristics of these types of systems in the current panorama.
  We want also to point out the significance of EFSs by studying all types of Data Mining problems in which they have shown a good behavior. In this way, we will stress the adapting capabilities of these systems by providing an exhaustive list of areas such as regression, classification, association rule mining, subgroup discovery and data streams, among others.
- Traditionally, the use of EAs over other optimization techniques has been considered as a natural choice due to their synergy with FRBSs. We seek to establish the reason behind this decision, understanding what the properties of EAs are, so that they make them to excel as opposed to other traditional approaches, i.e. neural networks.
- Since the beginning of EFSs, almost 25 years ago, there have been many real applications that stress that EFSs are still a valuable tool for solving different problems. Therefore, we will present a non-exhaustive list of several recent applications. Additionally, we will review some software tools that include EFSs. In this way, we may find algorithms of reference for contrasting any new developed method in this area.
- After this overview of the main aspects of current EFSs, we will focus on the central axis of this work, which is to set a discussion on the new trends and challenges for these types of systems. We will enumerate several novel problems and applications in which EFSs may achieve good results, especially those related to scalability and the scenario of Big Data [58,109], being defined as one hot topic that still needs to be addressed in detail. Furthermore, we will present additional topics and features that can help researchers focus on new possibilities for the continuous improvement of these techniques.
- Finally, we will carry out a critical evaluation on the design process for EFSs. In particular, we aim at establishing some guidelines for the definition of a systematic procedure in order to achieve the most correct development of EFSs.

We must point out that in this paper we do not aim at gathering an exhaustive list of references on the topic. We acknowledge that there has been an explosion of related works, which have already been partially covered in four previous reviews [36,38,53,75] and a book [39]. Within these works, we may find a profuse literature for EFSs, as well as in the following thematic Website at http://sci2s.ugr.es/gfs/.

The main contributions of the current paper with respect to the former references have been stressed in the fifth objective pointed out above. Specifically, we pay special attention to:

1. Extending the previous taxonomy on EFS [75] by proposing a new and more comprehensive one. As such, we have provided a more complete taxonomy by including those approaches that consider a trade-off between objectives (using MOEFs) and novel fuzzy representations.
2. Establishing why EAs are better suited for FRBS in contrast to other optimization techniques such as neural networks and ad hoc approaches
3. Presenting an overview and analysis of some of the most recent areas of Data Mining, such as classification with imbalanced data [101], mining of association rules [160], subgroup discovery [76], and low quality data [135], among others.
4. Enumerating some of the latest engineering applications with EFSs, in order to carry out a short overview for the significance of these types of techniques over real problems.
5. Introducing some software tools that allow researchers to work with EFSs, such as the KEEL Suite [11,9].
6. Developing an in-depth discussion for the new trends and challenges on the topic, thus analyzing hot and interesting areas for future work.

In order to face the aforementioned objectives, the remainder of this paper is organized as follows. In Section 2, we present the taxonomy on EFSs, and we describe the good properties that make EAs more suitable for FRBS over other optimization techniques. Section 3 reviews the use and behavior of EFSs over different types of Data Mining problems. Next, Section 4 introduces some recent real applications in which EFSs have shown to be a well-suited solution, whereas Section 5 describes the available software that includes these types of methods. Section 6 includes the main part of this manuscript in which we present and analyze several new trends and challenges for EFSs. Some brief remarks and guidelines for the development of new EFS are given in Section 7. Finally, in Section 8, we provide some concluding remarks.

2. Analyzing the Evolutionary Fuzzy Systems' models

The essential part of FRBSs is a set of IF-THEN fuzzy rules (traditionally linguistic values), whose antecedents and consequents are composed of fuzzy statements, related to with the dual concepts of fuzzy implication and the compositional rule of inference. Specifically, an FRBS is composed of a *knowledge base* (KB), that includes the information in the form of those IF-THEN fuzzy rules, i.e. the RB, and the correspondence of the fuzzy values, known as DB. It also comprises of an inference engine module that includes a fuzzification interface, an *inference system*, and a defuzzification interface.

EFSs are a family of approaches that are built on top of FRBSs, whose components are improved by means of an evolutionary learning/optimization process as depicted in Fig. 1. This process is designed for acting or tuning the elements of a fuzzy system in order to improve its behavior in a particular context. Traditionally, this was carried out by means of GAs, leading to the classical term of Genetic Fuzzy Systems [38,39,36,75]. In this paper, we consider a generalization of the former by the use of EAs [50].

Taking this into account, the first step in designing an EFS is to decide which parts of the fuzzy system are subjected to optimization by the EA coding scheme. Hence, EFS approaches can be

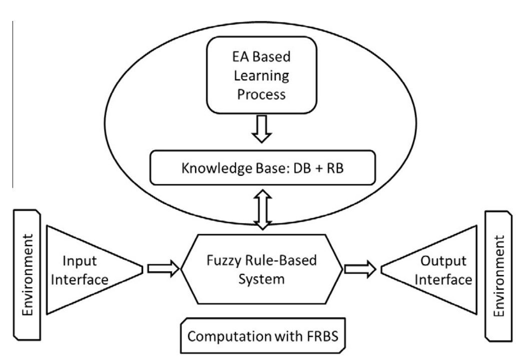

Fig. 1. Integration of an EFS on top of an FRBS.

mainly divided into two types of processes: tuning and learning. Additionally, we must make a decision whether to just improve the accuracy/precision of the FRBS or to achieve a tradeoff between accuracy and interpretability (and/or other possible objectives) by means of a MOEA. Finally, we must stress that new fuzzy set representations have been designed, which implies a new aspect to be evolved in order to take the highest advantage of this approach.

This high potential of EFSs implies the development of many different types of approaches. In accordance with the above, and considering the FRBSs' components involved in the genetic learning process, in this paper we extend the taxonomy first proposed by Herrera [75] by distinguishing among the learning of the FRBSs' elements, the EA components and tuning, and the management of the new fuzzy sets representation (see Fig. 2).

In order to describe this taxonomy tree of EFSs, this section is arranged as follows. First, we present those models according to the FRBS components involved in the evolutionary learning process (Section 2.1). Then, we focus on the multi-objective optimization (Section 2.2). Afterwards, we provide some brief remarks regarding the parametrized construction for new fuzzy representations (Section 2.3) Finally, we state the rationale for the selection of EAs approaches for carrying out the optimization process of FRBS (Section 2.4).

2.1. Evolutionary learning and tuning of FRBSs' components

When addressing a given Data Mining problem, the use of any fuzzy sets approach is usually considered when an interpretable system is sought, when the uncertainty involved in the data must be properly managed, or even when a dynamic model is under consideration. Then, we must make the decision on whether a simple FRBS is enough for the given requirements, or if a more sophisticated solution is needed, thus exchanging computational time for accuracy.

As introduced previously, this can be achieved either by designing approaches to learn the KB components, including an adaptive inference engine, or by starting from a given FRBS, developing approaches to tune the aforementioned components. Therefore, we may distinguish among the evolutionary KB learning, the evolutionary learning of KB components and inference engine parameters, and the evolutionary tuning. These approaches are described below, which can be performed via a standard mono-objective approach or a MOEA.

2.1.1. Evolutionary KB learning

The following four KB learning possibilities can be considered:

1. *Evolutionary rule selection.* In order to get rid of irrelevant, redundant, erroneous and/or conflictive rules in the RB, which perturb the FRBS performance, an optimized subset of fuzzy rules can be obtained [83].
2. *Simultaneous evolutionary learning of KB components.* Working in this way, there is possibility of generating better definitions of these components [77]. However, a larger search space is associated with this case, which makes the learning process more difficult and slow.
3. *Evolutionary rule learning.* Most of the approaches proposed to automatically learn the KB from numerical information have focused on the RB learning, using a predefined DB [149].
4. *Evolutionary DB learning.* A DB generation process allows the shape or the membership functions to be learnt, as well as other DB components such as the scaling functions, and the granularity of the fuzzy partitions. Two possibilities can be used: "a priori evolutionary DB learning" and "embedded evolutionary DB learning [40]."

2.1.2. Evolutionary learning of KB components and inference engine parameters

This area belongs to a hybrid model between adaptive inference engine and KB components learning. These type of approaches try to find high cooperation between the inference engine via parameters adaptation and the learning of KB components, including both in a simultaneous learning process [108].

2.1.3. Evolutionary tuning

With the aim of making the FRBS perform better, some approaches try to improve the preliminary DB definition or the inference engine parameters once the RB has been derived. The following three tuning possibilities can be considered (see the subtree under "evolutionary tuning").

1. *Evolutionary tuning of KB parameters.* A tuning process considering the whole KB obtained is used a posteriori to adjust the membership function parameters, i.e. the shapes of the linguistic terms [28].

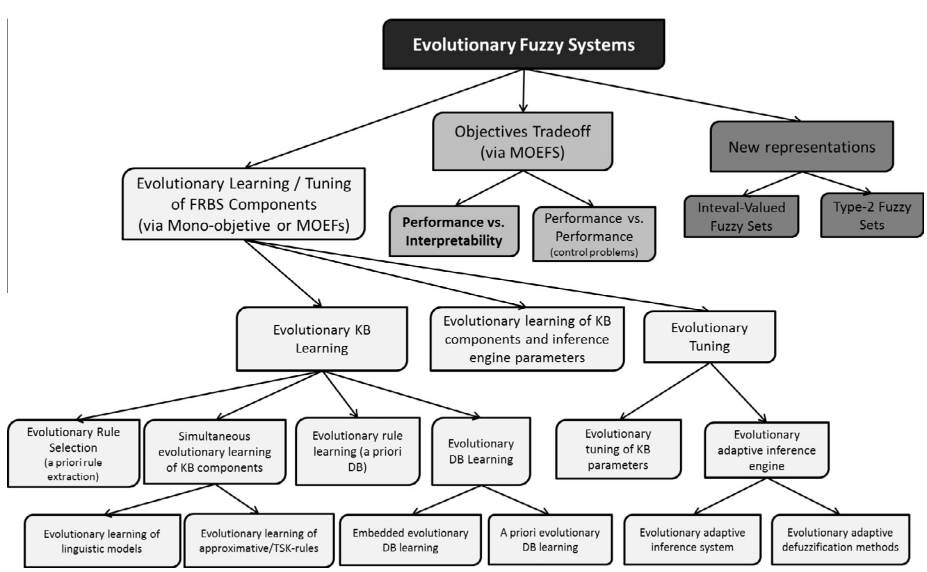

Fig. 2. Evolutionary Fuzzy Systems taxonomy.

2. *Evolutionary adaptive inference systems.* This approach uses parameterized expressions in the inference system, sometimes called adaptive inference systems, for getting higher cooperation among the fuzzy rules without losing the linguistic rule interpretability [10].
3. *Evolutionary adaptive defuzzification methods.* When the defuzzification function is applied by means of a weighted average operator, i.e. parameter based average functions, the use of EAs can allow us to adapt these defuzzification methods [91].

2.2. Approaches for optimizing several objectives

Traditionally, the efforts in developing EFSs were aimed at improving the accuracy/precision of the FRBS in a mono-objective way. However, in current applications the interest of researchers in obtaining more interpretable linguistic models has significantly grown [62]. The hitch is that accuracy and interpretability represent contradictory objectives. A compromise solution is to address this problem using MOEAs [32] leading to a set of fuzzy models with different tradeoffs between both objectives instead of a biased one. These hybrid approaches are known as MOEFSs [53] that, in addition to the two aforementioned goals, may include any other kind of objective, such as the complexity of the system, the cost, the computational time, and additional performance metrics.

In this case, the division of this type of techniques is first based on the multi-objective nature of the problem faced and second on the type of FRBS components optimized. Regarding the previous fact, those of the second level present a clear correspondence with the types previously described for EFSs in the previous section.

Here, we will only present a brief description for each category under consideration. For more detailed descriptions or an exhaustive list of contributions see [53] or its associated Webpage (http://sci2s.ugr.es/moefs-review/).

2.2.1. Accuracy-interpretability trade-offs

The comprehensibility of fuzzy models began to be integrated into the optimization process in the mid 1990s [81], thanks to the application of MOEAs to fuzzy systems. Nowadays, researchers agree on the need to consider two groups of interpretability measures, complexity-based and semantic-based ones. While the first group is related to the dimensionality of the system (simpler is better) the second one is related to the comprehensibility of the system (improving the semantics of the FRBS components). For a complete survey on interpretability measures for linguistic FRBSs see [62].

The differences between both accuracy and interpretability influence the optimization process, so that researchers usually include particular developments in the proposed MOEA making it able to handle this particular trade-off. An example can be seen in [61] where authors specifically force the search to focus on the most accurate solutions.

2.2.2. Performance vs. performance (control problems)

In control system design, there are often multiple objectives to be considered, i.e. time constraints, robustness and stability requirements, comprehensibility, and the compactness of the obtained controller. This fact has led to the application of MOEAs in the design of Fuzzy Logic Controllers.

The design of these systems is defined as the obtaining of a structure for the controller and the corresponding numerical parameters. In a general sense, they fit with the tuning and learning presented for EFSs in the previous section. In most cases, the proposal deals with the postprocessing of Fuzzy Logic Controller parameters, since it is the simplest approach and requires a reduced search space.

2.3. Novel fuzzy representations

Classical approaches on FRBSs make use of standard fuzzy sets [159], but in the specialized literature we found extensions to this approach with aim to better represent the uncertainty inherent to fuzzy logic. Among them, we stress Type-2 fuzzy sets [90] and Interval-Valued Fuzzy Sets (IVFSs) [127] as two of the main exponents of new fuzzy representations.

Type-2 fuzzy sets reduce the amount of uncertainty in a system because this logic offers better capabilities to handle linguistic uncertainties by modeling vagueness and unreliability of information. In order to obtain a type-2 membership function, we start from the type-1 standard definition, and then we blur it to the left and to the right. In this case, for a specific value, the membership function, takes on different values, which are not all weighted the same. Therefore, we can assign membership grades to all of those points.

For IVFS [127], the membership degree of each element to the set is given by a closed sub-interval of the interval $[0,1]$. In such a way, this amplitude will represent the lack of knowledge of the expert for giving an exact numerical value for the membership. We must point out that IVFSs are a particular case of type-2 fuzzy sets, having a zero membership out of the ranges of the interval.

In neither case, there is a general design strategy for finding the optimal fuzzy models. In accordance with the former, EAs have been used to find the appropriate parameter values and structure of these fuzzy systems.

In the case of type-2 fuzzy models, EFSs can be classified into two categories [29]: (1) the first category assumes that an "optimal" type-1 fuzzy model has already been designed, and afterwards a type-2 fuzzy model is constructed through some sound augmentation of the existing model [30]; (2) the second class of design methods is concerned with the construction of the type-2 fuzzy model directly from experimental data [154].

Regarding IVFS, current works initialize type-1 fuzzy sets as those defined homogeneously over the input space. Then, the upper and lower bounds of the interval for each fuzzy set are learnt by means of a weak-ignorance function (amplitude tuning) [139], which may also involve a lateral adjustment for the better contextualization of the fuzzy variables [138]. Finally, in [140] IVFS are built ad hoc, using an interval-valued restricted equivalence functions within a new interval-valued fuzzy reasoning method. The parameters of these equivalence functions per variable are learnt by means of an EA, which is also combined with rule selection in order to decrease the complexity of the system.

2.4. The rationale for Evolutionary Fuzzy Systems

As previously mentioned in this manuscript, EAs are a general technique that can be used for obtaining FRBSs in a natural way. Unlike the ad hoc approaches, EAs provide a generic code structure and independent performance features with the flexibility and capability to incorporate existing knowledge. This a priori knowledge in an FRBS learning process can be given in several ways, i.e. relevant inputs, scaling factors, membership functions, shape functions, granularity, fuzzy rules, inference parameters, number of rules, and so on. All of them can be represented in the genetic code and/or considered in the genetic search by defining different mechanisms for managing them, such as genetic operations, niches, and coevolution among others. In addition, as discussed above the EAs ability to explore large search spaces for suitable solutions only requiring a performance measure can be used to face the FRBS design with a wide range of approaches: RB learning with a predefined DB, DB learning, RB and DB learning, rule selection, KB, inference system or defuzzification methods tuning.

These characteristics are very valuable for the design of FRBSs and bring benefits to EFSs over adhoc approaches.

Other Computational Intelligence techniques as neural networks have been applied to the design of FRBSs, the so-called neuro-fuzzy systems (NFSs). These types of systems is characterized by a fuzzy system where fuzzy sets and fuzzy rules are adjusted using the structure and learning capabilities of neural networks. Nevertheless, in most of the NFSs proposals as FALCON [96], GARIC [21], ANFIS [126], FINEST [148], NEFCON [114] and SONFIN [88], the learning process only concerns the adaptation of internal parameters of a fixed structure. This fact implies some limitations for NFSs that EFSs overcome. NFSs suffer the curse of dimensionality in a higher degree due to the complexity of the geometrically increase of the neural learning process with respect to the number of variables. Usually, NFSs need to know the granularity (number of labels considered) prior to the learning process. On the contrary, EFSs may include the granularity as another parameter in the evolutionary process. Furthermore, NFSs have difficulties for learning the rule structure because usually NFSs only determines membership functions and rule consequent coefficients. As mentioned above, EFSs can face the full learning of KB, and evolutionary approaches as genetic programming make the learning easier inside the rule structure.

Finally, every Intelligent Data Analysis process must optimize different objectives such as accuracy, interpretability, cost, computational time, additional performance metrics, just referring to some of them. With the use of MOEFSs, users can take into account multiple goals within the optimization process, hence generating a set of non-dominated fuzzy systems that represents a tradeoff among objectives. Just like other EFSs, MOEFSs have been applied in several fields, due to their ability to represent real-world problems in a simple way and to include a priori knowledge in the model.

3. The success of EFSs on Intelligent Data Analysis

One of the key values that has enabled EFSs to become an important tool in computational intelligence is clearly its good results and robustness in practically all fields in Data Mining. In this section, we carry out an overview of some of these problems, pointing out the most significant contributions in each topic.

3.1. Classification and regression problems

EFSs have been profusely used in the classification and regression framework. In previous reviews on the topic [38,53,75] we may find them widely mentioned. Therefore, we will give a summary of this subject in order to focus on less developed topics, but comprise of a higher number of contributions nowadays.

- The assumptions that bind standard regression analysis can be relaxed by using a fuzzy regression model. Indeed, the use of fuzzy sets has paved the way for the improvement of modeling tasks, particularly by the use of the Takagi Sugeno Kang (TSK) system [147]. Therefore, the extension to EFSs was a natural progression. In particular, most of the developments for learning and tuning described in Section 2.1 were first proposed for regression and control problems. Recent developments using MOEFSs can be found in [63], and high dimensional problems via adaptive fuzzy inference systems [107].
- Classification is one of the most studied problems in Machine Learning and Data Mining [74]. It is a task that, from a supervised learning point of view, consists of inducing a mapping which allows to determine the class of a new pattern from a set of attributes. An algorithm is used to generate a classifier from a set of correctly classified patterns called training set. In this framework, if we join the use of fuzzy sets to the design of rule based systems, we will obtain what is known as Fuzzy Rule Based Classification Systems (FRBCSs) [82].
The actual number of research works that cover this topic is overwhelming. Therefore, it is impossible to excel a single set of publications that represent this area.

3.2. Classification with imbalanced data

When working with real applications in classification, we can see that they frequently present a very different distribution of examples inside their classes, which is known as the problem of imbalanced classes [101]. Linguistic FRBSs have shown the achievement of a good performance in the context of classification with imbalanced datasets [59]. Specifically, linguistic fuzzy sets allow the smoothing of the borderline areas in the inference process, which is also a desirable behavior in the scenario of overlapping, which is known to highly degrade the performance in this context [70].

Solutions in imbalanced classification, are usually divided into three different groups as follows:

- Data-level approaches for rebalancing the training set, where the most widely used technique is the SMOTE algorithm [31]. In this group we may find most of the examples of EFS algorithms [56,57,153].
- Algorithmic approaches are designed with specific operations for the skewed class distribution. Within this group we must stress FLAGID [144], the use of a hierarchical FRBCS [55,100], and the use of a feature weighting FRBCS for addressing the overlapping areas [12]. Additionally, other approaches have been designed ad hoc for addressing the imbalance problem in Intrusion Detection Systems [51], and the modeling and prediction of imbalanced financial datasets [137].
- Cost-sensitive learning approaches have also shown a behavior countable in this context. EFSs solutions can be found using standard fuzzy learning algorithms [122], an MOEFS [49], or an ensemble method [145].

3.3. Mining of association rules

Association rules are used to represent and identify dependencies between items in a database [160]. These are an expression of the type $X \to Y$, where X and Y are sets of items and $X \cap Y = \emptyset$. This means that if all the items in X exist in a transaction then all the items in Y with a high probability are also in the transaction, and X and Y should not have any common items [3]. Knowledge of this type of relationship can enable proactive decision making to proceed from the inferred data. Many problem domains have a need for this type of analysis, including risk management, medical diagnostics, fire management in national forests and so on.

The first studies on the topic focused on databases with binary values, however the data in real-world applications usually consist of quantitative values. Therefore, the use of fuzzy sets to describe association between data extends the types of relationships that may be represented, facilitates the interpretation of rules in linguistic terms, and avoids unnatural boundaries in the partitioning of the attribute domains [48]. Whereas classical algorithms use a predefined DB, most recent approaches on fuzzy association rule mining are focused to learn both the fuzzy rules and the membership functions of the fuzzy labels [7,78].

3.4. Subgroup discovery

Subgroup discovery [76,94] is a form of supervised inductive learning of subgroup descriptions in which, given a set of data and having a property of interest to the user, attempts to locate subgroups which are statistically "most interesting" for the user. Its objective is to obtain simple rules (with an understandable structure and with few variables), that are highly significant and highly supported (covering many of the instances of the target class).

A fuzzy rule describing a subgroup is represented in the same way as for classification tasks. The antecedent describes the subgroup in either canonical or disjunctive normal form, and the consequent represents a value for the target classes. Currently, some approaches make use of fuzzy logic for representing the continuous variables that form the antecedent of these rules, by means of linguistic variables. Since subgroup discovery requires several objectives, we may excel MOEFSs approaches such as NMEEF-SD [25]. In addition, several applications for subgroup discovery have been developed in areas like marketing [44], photovoltaic technology [27], or medical diagnosis [24].

3.5. Learning from low quality data

There are many practical problems requiring learning models from uncertain data. The experimental design for the EFS learning from data observed in an imprecise way, is not being actively studied by researchers. However, according to the point of view of fuzzy statistics, the primary use of fuzzy sets in classification and modeling problems is for the treatment of vague data. Preliminary results in this area involve the proposals of different formalizations for the definition of fuzzy classifiers, based on the relationships between random sets and fuzzy sets [134].

Throughout the years, several works have proposed the use of vague data with EFS under different perspectives [121,135]. Furthermore, these approaches has been used in several applications such as athletics [120] and diagnosis of dyslexia [119].

3.6. Data streams

In many applications, developed systems must face a dynamic environment which implies the use of an adaptive learning approach, knowing as learning from data streams [64]. Specifically, the key motivation for this task is extracting knowledge structures from an ordered sequence of records/instances that arrive to the system in a rapid way, such as in computer network traffic or sensor data. Therefore, the system learns incrementally, and maybe even in real-time, aiming to adapt itself to changes of environmental conditions or properties of the data-generating process.

Regarding EFSs, just a few proposals have been developed due to the time constraint requirements for this scenario. The first approach was carried out in [116], whereas two applications can be found at [150] for modeling real state market and in [142] for monitoring rolling mills.

4. Recent applications of EFSs

As EFSs have evolved to better address the imprecision and uncertainty in data, they have further attracted the interest of practitioners to address specific problems. The optimization obtained by EAs is a powerful boost in the performance of the proposed approaches and also in the interpretability of the obtained designs. The significance of these types of models is supported by the high amount of real application problems in which they are applied.

Table 1 shows a short list of some recent applications that have been addressed using EFSs. This list is organized according to the publication year of the associated research work, so that the applications can be consulted from old to new.

In a quick glance, we can observe that the applications shown are quite varied and that they are related to diverse general subjects. Specifically, we can observe applications in computing, medicine, finance, healthcare, industry and so on. This demonstrates the flexibility of EFSs, as they are able to handle problems that at least initially are not especially similar. Moreover, the number of research publications containing applications of EFSs has increased in the last few years, which is also a clue about the importance of these methods in real-world problems and the attraction that they are still generating.

5. Software suites for using EFSs: KEEL and frbs R package

As more and more EFS approaches are being developed over the years, there is also a growing interest in compiling all these new algorithms into a single library according to a twofold necessity: (1) being able to compare our novel proposals versus the state-of-the-art on the topic; and (2) to analyze and carry out a thorough experimentation for determining the best suited solution for a given application problem.

KEEL (Knowledge Extraction based on Evolutionary Learning) Software Suite[1] [11,9] is a free (GPLv3) Java tool which empowers the user to assess the behavior of evolutionary learning and soft computing based techniques for different kind of Data Mining problems: regression, classification, clustering, pattern mining and so on. Particularly, it supports a complete list of Fuzzy Rule Based Learning algorithms, both standard and evolutionary, together with methods for performing an evolutionary postprocess phase over the results of a Fuzzy Rule extraction method (only for regression tasks).

The significance of the KEEL Software Suite is mainly based on the aforementioned exhaustive set of EFS algorithms that are included. In particular, it includes some recent approaches such as the different versions of SLAVE [65], FARC-HD [8] and its extension to IVFS [140], as well as well-known and classical algorithms such as the MOGUL methodology [37], the Fuzzy-Hybrid Genetics Based Machine Learning approach [84], and the genetic lateral tuning based on 2-tuples [5], among others.

We may also find an R package named as *frbs*[2] which provides over ten learning methods to construct FRBS models in regression and classification tasks from available data, and also allows the user to work with several model types, i.e. linguistic, TSK, and so on. Regarding EFS, it implements some of the former algorithms already enumerated for the KEEL Software Suite, thus stressing the completeness of this R library [141].

6. New trends and challenges for EFSs

We must also acknowledge that, in spite of the good behavior shown in many Data Mining areas, and their success when they have been applied on different real problems, the behavior of EFS can still be improved if we take advantage of the novel features and components that have been recently proposed for the evolutionary models, in terms of both performance and scalability.

In this section, we will present a discussion on the new trends and challenges for EFSs. These issues must be regarded as a path for future work and thus, they must allow researchers to focus on the most prolific topics for developing their new methodologies. In order to do so, we will first introduce some original types of EAs

[1] http://www.keel.es.
[2] http://CRAN.R-project.org/package=frbs.

Table 1
Recent applications of EFSs.

Year	Applications of EFSs
2011	Computing, intrusion detection in computer networks [1,152]
	Medical, characterization of patients in the psychiatric emergency [26]
	Chemical, identification of virus to filter the water microfiltration [105]
	Sports, future performance modeling in athletism [120]
2012	Healthcare, modeling the human gait [13]
	Financial, time series forecasting of banking products [17]
	Anthropology, skull forensic identification [80]
	Financial, stock price prediction [97]
	Musical, classifying audio signals in sound types [143]
2013	Medical, analysing the impact of a virus [24]
	Energy, exception identification in photovoltaic technology [27]
	Healthcare, dynamic activities recognition using an accelerometer [52]
	Remote sensing images [113]
	Marketing, academic research and management support in marketing [117,118]
	Industrial, characterizing rotor blade learning edge materials [136]
	Financial, modeling the real estate markets [150]
2014	Industrial, intelligent data analysis in industry [16]
	Computing, intrusion detection [51]
	Medical, categorizing risks for the coronary heart disease [47]
	Medical, assessing the mortality risk after a cardiac surgery [115]
	Financial, modeling and predicting real-world financial applications [137]
	Industrial, fault detection at rolling mills [142]
	Web recommendation systems [132]
	Ecology, river flows forecasting [151]

(Section 6.1) and the scalability from the point of view of EAs (Section 6.2). Next, we will point out how an instance selection preprocessing stage can be carried out with the help of EFS (Section 6.3). Then, we will present the problem of Big Data, as being one of the hottest topics for current research (Section 6.4). New and complex classification problems will be described next (Section 6.5). Finally, we will consider the benefits of addressing intrinsic data characteristics of the datasets (Section 6.6).

6.1. On the use of different types of EAs

EAs are certainly the basis for the development of EFSs. From traditional GA with binary representations, to more sophisticated real-coded evolutionary approaches, researchers are taking advantage of the learning/optimization capabilities that arise with every new approach that is being released. Regarding this fact, we must pay attention to the recent developments in this field of research, and how they can be employed to improve the behavior of EFSs.

A common example is the use of MOEAs [32] in the design of MOEFs [53], as suggested in Section 2.2. The benefit in this case is twofold: (1) to perform the search optimizing several objectives at once; and (2) to spread the space of solutions and to be able to select and/or combine the final set of learned models. It is also worth exploring the advantages that new approaches such as NSGA-III [43,85] may provide, as well as analyzing the use of MOEAs for Data Mining [112] and to extend them to FRBS.

We must also mention the use of niching-based multimodal algorithms [124]. This comprises a similar case to the former in which, according to the intrinsic drawback of standard EAs of converging to a population containing just one solution, we preserve genetic diversity by encouraging the formation of species or niches, each representing one of the possible solutions. In this way, the search abilities of the algorithm are enhanced to get different solutions in a multi-modal problem.

Finally, micro-GAs [2,33], which are based on small populations, have shown to be efficient as they can converge in a lower number of iterations. This does not mean that an optimal solution has been attained, but it gives the possibility of restarting the search by using several configuration sets, i.e. the combination of genetic operators and/or their values, for obtaining a better global system. Therefore, these algorithms may be useful for the optimization of complex models.

6.2. Scalability from the point of view of EAs

The scaling up problem has arisen as a very influential one in Machine Learning and Data Mining [35]. Essentially, the scalability defines the capability of an algorithm for maintain a proper level of performance, at a reasonable computational cost, as the size of the problem increases. Specifically, we may distinguish between two kinds of problems that should be solved by applying different kinds of techniques but both together: (1) high dimensional problems when a large number of variables have to be considered and, (2) large scale datasets, when managing problems comprised of a large amount of data.

Both types of problems have a strong influence in linguistic fuzzy modeling. On the first hand, because the learning capability of FRBSs suffer from exponential rule explosion when the number of variables and/or data examples becomes high [34,87]. On the other hand, because of the already mentioned increase of the training time, but also the convergence towards compact and interpretable models [75].

The hitch here is that EAs are computationally expensive by themselves. Usually, a large number of evaluations are needed to reach convergence, and depending on the size of the problem, each call to the fitness function may take a long time. For Data Mining tasks in general, and for EFS in particular, there are several techniques available that improve the performance of the traditional algorithms in these situations [71]. Most of them can be classified into three perspectives:

1. **Algorithm oriented** [125]. It includes those techniques that modify the definition of the algorithm adapting its components, or integrating any mechanism to improve its performance when tackling large scale data [6,20,63]. This has been achieved by means of the design of algorithms that integrate any mechanism oriented towards the problem dimensionality.
2. **Data oriented** [69]. These types of techniques are applied directly to the data, reducing and distributing it with the objective of reducing the computational cost of the traditional Machine Learning techniques. This can be carried out by means of feature selection processes [18], instances or prototype selection [72], or by modifying the example distribution [42]. These preprocessing approaches can be carried out as part of the objective when we aim at improving the quality of the data before the training stage.
3. **Distributed approaches** [4]. In this case, the methodology is developed to run on several machines in order to save computational time [129].

Whereas the first type of solution requires a thorough design of the inner features of the EA to achieve better efficiency, the latter two types can be viewed as "external approaches" that are somehow independent of the learning stage of the EFS. Therefore, authors must be encouraged to focus on new methodologies for decreasing the time complexity of this type of algorithms.

6.3. Instance selection by means of EFSs

It is well-known than one of the main drawbacks of EAs and consequently for EFSs, is the computational time consumed during the fitness evaluation. Clearly, the size of the training set has a strong dependency with the former fact, but it can also affect the complexity of a resulting model. This is due to the fact that a higher number of instances generally induces the generation of FRBS with a higher number of rules. The way to overcome both problems, at least partially, is to carry out a preprocessing stage by means of a training set selection method. The goal is trying to maintain the global accuracy of the model by using a representative set of examples from the initial dataset.

In [54], authors presented an analysis of the influence of instance selection methods for EFS, by considering 36 different methods from the state-of-the-art. They have determined that the instance selection technique depends basically on the dimension of the considered dataset. A quite different approach can be found in [15], where authors have embedded the instance selection stage within the whole EFS process, leading to a co-evolutionary approach. Specifically, they made use of a MOEFS for learning the FRBS, and a standard GA for obtaining a reduced training set. They show that with just a set between the 10% and 20% of the original instances, the Pareto front is comparable to the one discovered with the full dataset, but reduces the running time up to 86%.

In accordance with the conclusions extracted in these papers, further work should be carried out following this research line. Researchers must focus on the development of a methodology that allows a fine-tuned integration between the instance selection and the EFS learning approach. Proceeding this way, new techniques can be better suited for scalability purposes, as suggested in Section 6.2.

6.4. Big Data

Recently, the term of Big Data has been coined referring to those challenges and advantages derived from collecting and processing vast amounts of data [58]. This topic is mainly identified by the 3V's definition which includes a large *volume* of information, which arrives at a high *velocity* rate and maybe with real-time requirements, and with a *variety* of structured, semi-structured or unstructured representation. Additional definitions including up to 9V's can be also found, adding terms like Veracity, Value, Viability, and Visualization, among others [164].

There are several tools that have been designed to address Big Data problems. Among them, the most popular one is the *MapReduce* distributed processing system, as well as its open source counterpart *Hadoop* [93]. These models allow the system to automatically parallelize the execution, which can be achieved by the simple definition of two functions, referred to as Map and Reduce. In short, "Map" is used for per-record computation, whereas "Reduce" aggregates the output from the Map functions and applies a given function for obtaining the final results. In addition to these systems, *Spark* [89] is a new system developed to overcome data reuse across multiple computations. It supports iterative applications, while retaining the scalability and fault tolerance of MapReduce, supporting in-memory processes. This latter emergent technology aspires to be the new reference for the processing of massive data.

New research in the Big Data scenario for EFSs must be focused on adapting the state-of-the-art algorithms, i.e. the ones presented in Section 5, by providing a thorough implementation of the Map and Reduce functions with aims at taking the highest advantage of the parallelization for both reducing the learning time costs and maintaining the overall accuracy. We must stress that this adaptation is not trivial, and requires some effort when determining how to proceed in a divide and conquer method from the original workflow.

Being a recent framework, there are just a few works which address this topic (using the MapReduce paradigm) from the point of view of EAs [95], and even fewer from the perspective of FRBSs [99,98]. This is due to the fact that MapReduce has a significant performance penalty when executing *iterative jobs* [58], as it reloads the input data from the disk every time, regardless of how much it has changed from the previous iterations.

In accordance with the above, the Spark programming framework must be considered as a most productive tool that may lead to more efficient implementations for the field of EFSs. Furthermore, we have stressed that research with FRBSs and EFSs is on an initial stage and only a couple of first approaches on the topic have been published. There is still a lot of work to be carried out in this sense.

6.5. Complex classification problems

The classification problem can been addressed under different perspectives depending on the properties of the problem, and the final objective that is pursued. Among different possibilities, we have stressed three problems that currently hold high interest in the research community and that have not be widely addressed with EFSs yet. Specifically, in this section we describe multi-label, multi-instance and monotonic classification.

6.5.1. Multi-label classification

In conventional classification, each instance is assumed to belong to exactly one class among a set of candidate types. However there is a framework called "multi-label classification" whose goal is to obtain a model that classifies a set of non-exclusive categories or, in other words, to assign more than one tag the same example [161].

The multi-label classification has received increased attention in recent years according to its practical relevance, and its interest from a theoretical point of view. This type of classification task was first applied in the field of document categorization, but it is currently applied to multiple areas of interest including computer vision, information extraction through acoustic signals, and bioinformatics, among others.

The nature of the problem is well suited for the use of FRBS and EFS based algorithms, but few efforts have been applied to solving this task using these tools. We may find just a couple of contributions regarding this issue, either using a fuzzy similarity measure and k-nearest neighbors [86] and a fuzzy associative classifier with genetic rule selection [146].

By means of the fuzzy rule representation, we may extract and represent the features of this complex problem in an accurate way. In addition to the former, and according to the few works that currently cover this topic with EFS, we must be aware of the necessity in filling the gap in this field of research.

6.5.2. Multi-instance learning

Multi-instance learning (MIL) [46] is a variation on the classic paradigm supervised learning problems in which we have incomplete information about classes or instances. Unlike the classical model, where instances are categorized individually, in MIL "bags" of instances are classified as a whole. However, it is unknown which one of these instances is the one that categorizes the full "bag". This is a field with many practical applications, where the number and variety of these areas that have benefited from MIL has increased significantly. We stress bioinformatics, Web recommendation, and computer-aided detection, among others.

There are dozens of methods that search for a solution and new proposals to improve specific aspects emerge continuously [14]. The use of fuzzy related models, may help to consider lower and upper approximations of the set of bags, using similarity measures in this level. We found a first approach that applies FRBS [106]. The extension to EFS may unequivocally improve any preliminary results achieved by these techniques. Finally, we must point out an additional open problem, which consists of the hybridization between multi-label and multi-instance learning [162], and that might be also addressed in the near future.

6.5.3. Ordinal and monotonic classification

Ordinal reasoning is associated with an important category in prediction problems. In the ordinal classification task, the output attribute takes ordinal values within certain domains [23], so that accuracy is computed by measuring how "far" the class is from the actual one.

In addition to the former, we may observe some problems in which both the input attributes and the class have a monotonicity constraint. This is known as monotonic classification [92], where there partial ordering relationships exist that establish the order among examples, being in particular a case of the former.

The hitch in this classification scenario is that those models obtained by standard techniques do not ensure compliance with the main monotonicity restriction previously pointed out. This has motivated the development of ad hoc algorithms that are able to handle such restrictions, in particular with fuzzy rule based similarity [19]. Furthermore, solutions based on data processing are beginning to be studied, for example those based on fuzzy ordinal measures for feature selection [79].

EFSs may help to address this problem from an algorithmic point of view. This can be carried out by means of the development of ad hoc techniques that take advantage of the fuzzy representation for a better fitting of the output ranking, as well as providing the proper metrics to guide the whole optimization process [41].

6.6. Addressing data intrinsic characteristics: overlapping, small disjuncts, noise and dataset shift

It is straightforward to acknowledge that there are some problems that are considered to be harder, or more difficult to solve accurately, than others. The hitch here is being able to identify the properties of these problems prior to the learning stage, and therefore to apply specific methodologies for improving the behavior of standard Data Mining approaches.

A recent study made in [101] presented a discussion about six significant problems related to data intrinsic characteristics that, in addition to the well-known class imbalance (refer to Section 3.2), contribute to a performance degradation. These are the problem of overlapping between the classes [133], the identification of areas with small disjuncts [157] together with the lack of density and information in the training data [155], the impact of noisy data [163], and the dataset shift [110]:

- Among all data intrinsic characteristics, we must highlight the one related to overlapping between classes [133], which appears when a region of the data space contains a similar quantity of training data from each class. This situation leads to developing an inference with almost the same a priori probabilities in this overlapping area, which makes the distinction between the classes very hard or even impossible. Indeed, some researchers have shown that the behavior of a given classifier over a dataset can be characterized by means of some data complexity metrics that measure the overlap between the classes [104]. Taking this into account, we may find a recent work with EFSs that overcomes this issue, in which authors proposed a simple yet effective feature weighting approach in synergy with FRBCS for imbalanced classification [12].
- The presence of imbalanced classes is closely related to the problem of small disjuncts. This situation occurs when the concepts are represented within small clusters, which arise as a direct result of underrepresented subconcepts [158]. We must we aware that this problem becomes accentuated for those classification algorithms which are based on a divide-and-conquer approach [156], leading to data fragmentation. Somehow similar is the small sample size [128]. This issue is related to the "lack of density" or "lack of information" where induction algorithms do not have enough data to make generalizations about the distribution of samples. We must stress that FRBSs have never been proposed to deal with any of these cases yet.
- Noisy data is known to affect the way any Data Mining system behaves [60]. The most common approach is to apply a filtering technique in order to avoid the influence of examples with noise in the learning stage [131]. Although fuzzy set representation is known to handle uncertainty in an adequate way, and they have shown to support a good tolerance versus noise [130], to the best of our knowledge there are no significant contributions that have made use of these types of models for managing noisy data in classification problems. In accordance with the former, we must stress this framework for carrying out new research on the topic.
- The problem of dataset shift [110,22] is defined as the case where training and test data follow different distributions. It often appears due to sample selection bias issues, being covariate shift the most common case, i.e the input attribute values that have different distributions between the training and test sets [110]. Therefore, a more suitable validation technique has been designed in order to avoid introducing dataset shift issues artificially [111,103], showing a robust behavior in the case of EFS [102].

In summary, we have stressed the significance of identifying the inner characteristics of the data as a prior step when studying the applications we are dealing with. Proceeding this way, the benefit is straightforward, as we may use or develop solutions that are better suited to each case study. Clearly, the learning capabilities of fuzzy systems, and the good search properties and adaptability of EAs, make EFSs a suitable methodology for proposing solutions that are able to properly manage Data Mining problems with respect to these properties and different complexity measures of the Data Mining problems.

7. Methodology for the development of EFSs

In this section we aim to carry out a critical discussion on the procedure followed by researchers for the design and development of EFS. In particular, we will focus on two different aspects that must be considered in order to advance towards a correct methodology and strengthen the progress in this framework, i.e. the inner design structure of EFS, and the correct experimental evaluation for novel proposals.

7.1. Alternatives for the EFS structure

There are two main choices that must be taken into account when developing a new EFS, i.e. the components of the FRBS (including those elements that will be learnt or tuned by the EA), and the EA scheme and inner features for developing the optimization. We describe both below.

1. **The definition of the fuzzy system**. Researchers must traditionally decide on several design parameters such as universes granulation, rule antecedent aggregation operators, rule semantics, rule base aggregation operators and defuzzification methods, among others. This implies a high number of degrees of freedom that must be studied thoroughly.

   The most common approach is to define a standard KB and, depending on the problem requirements, to refine any of its parameters. Some examples in this case are the use of novel fuzzy representations, the definition of approximative rules (instead of descriptive ones), using parametric connectors, rule weight heuristics, and so on. Researchers may take advantage of the EAs for encoding and evolving some of these components, and optimize FRBSs with respect to the design decisions above. However, they must be careful as a very ambitious approach will increase the chromosome complexity, enlarging the search space considerably. This implies a challenge on the computational efficiency of EA-based methods, that, as pointed out previously, can be attenuated only via judicious exploration of design requirements.

2. **The components of the evolutionary part**. When working with EAs, we must be aware of useful elements such as real coding for continuous variables, different parent replacement strategies, and adaptive components. This supposes a step forward with respect to the simplest GA that researchers have traditionally been used even for new approaches.
   Additionally, recent literature on EAs has introduced significant advances such as novel techniques (particle swarm, differential evolution, and so on), and the extension of standard GA components such as niching GAs for multimodal functions, hybrid combinations of GAs and local search (memetic algorithms). The choice of a more sophisticated EA may result in better learning and adaptation and therefore to a more accurate EFS.
   However, when these new types of techniques are being employed for the learning and/or tuning of the FRBS components, a clear justification for their choice must be made from whatever meaningful point of view: efficiency, efficacy/precision, interpretability, scalability, and so on. In summary, a comparison with the classic EFSs from the literature must be carried out in order to establish whether these new features actually improve the proposed system.

7.2. Proper experimental validation for EFSs

Focusing on the experimental study, there are four different issues that must be analyzed, namely the comparison with the state-of-the-art approaches, proper use of experimental framework (benchmark problems, validation techniques, etc.), statistical analysis for the support of experimental results, and reproducibility of the proposal.

1. **Comparison with the state-of-the-art**. When any new approach is under study, the scientific method implies the justification of its usefulness and quality with respect to any metric of performance, i.e. accuracy, complexity, and so on. It is mandatory to carry out a experimental analysis by contrasting the results of the authors' proposal with the previous best techniques which follow the same objectives and fuzzy system components, following the branches of the EFS' taxonomy shown in Section 2. Proceeding this way, the effectiveness of the novel algorithm for this knowledge area will be unequivocally shown.
   Unfortunately, there is no review study that establishes which are the specific algorithms that must be stressed for each single category, and even if this was the case, it is clear that new approaches are published every year which outperform the previous ones. Nevertheless, authors must follow two simple recommendations: (1) to compare with the most well-known and classical approaches in the literature and (2) to determine the advantages with respect to the most recent proposed techniques in the topic and that present high quality results. It is not enough to compare with a simple approach that was already outperformed years ago, that is not the state of the art at the present, or with novel methods that do not accomplish the aforementioned requirements.
2. **Setting up a correct experimental framework**. Selecting the proper algorithms for comparison (as pointed out above) is very important, but if the framework used in the experimental study is incorrect, the findings extracted from the results may no longer be valid. Every paper that is being published usually employs a different set of benchmark problems, or it is based on a specific application from which it is impossible to reproduce the same experimental study, or even based on simple toy problems created ad hoc with synthetic samples.
   The previous issue implies the management of a unified set of datasets for different types of Data Mining problems (please refer to Section 3). Furthermore, it is also indispensable to make use of the same data partitions and validation techniques. In this sense, we must stress the development of the KEEL dataset repository [9] that can be found at http://www.keel.es/datasets.php.
   In addition to the former, we must point out the necessity of using experimental analysis models under an equal and significant number of runs, iterations, parameters, and execution time.
3. **Statistical tests**. Traditionally, authors claim that their new proposals are better than the algorithms of comparison by just focusing on the absolute difference shown in the selected metrics of performance. However, this practice is no longer correct as it is mandatory to establish some statistical support to these findings. Inferential statistics show how well a sample of results supports a certain hypothesis and whether the conclusions achieved can be generalized beyond what was tested.
   There are two main types of statistical tests in the literature: parametric tests (*t*-test and ANOVA) when the results fulfill the independency, normality and homoscedasticity conditions [66]; and nonparametric tests (Wilcoxon or Friedman tests) when the former assumptions do not hold. In EFS and other computational intelligence algorithms, these conditions are not easy to meet [67].
   In accordance with the above, we suggest the use of non-parametric tests as stated in the specialized literature [45,68]. The use of pairwise or multiple-tests depends on whether we are contrasting our control method versus a single algorithm or versus several related algorithms. In summary, statistical tests are a necessary tool to strengthen the conclusions extracted from the experimental results. They allow significant differences to be shown among the results of the approaches that are being compared, highlighting the best performing method.
4. **Reproducibility of the new algorithm**. Finally, it is of extreme importance to be able to reproduce the results of the authors' proposal according to a double aim: (1) to confirm the results shown in the experimental study and (2) to carry out a future comparison when developing a related approach.
   A clear algorithmic description of the proposal must be carried out, including all components (coding approach, operators, fuzzy components, etc.). All steps should be clearly identified, with no possibility of taking ad hoc choices by other researchers at any point. Additionally, all parameters' values must be explicitly pointed out in the experimental framework.
   In order to ease this issue, a good exercise is to include the new approach into any Data Mining tool such as those introduced in Section 5, or at least to make the source code publicly available for other users.

8. Concluding remarks

In this survey paper we have carried out a full review of EFS models. First, we have presented a complete taxonomy for the current types of associated methodologies with respect to three different perspectives, i.e. the learning of the FRBS' elements, the approaches with respect to the evolutionary components, and the optimization of novel fuzzy representations. Additionally, we have stressed the significance of EFS regarding the good behavior shown over many different Data Mining scenarios, and also their application for solving real problems in heterogenous fields. With the sake of complementing the overview of EFSs, we have also pointed the software suites that include these types of models.

In accordance with the former, we have analyzed and evaluated the good properties and features of EFSs, and their success for traditional frameworks. However, in this paper we aimed to go one step further by introducing some topics that we considered to be as new trends and challenges for this area. Specifically, we have described several fields of Data Mining in which EFS still have room for improvement, and also those novel work scenarios that need the development of more sophisticated approaches. Among them, we have stressed new and complex classification problems, the scalability problem from the point of view of EAs and also its relationship with Big Data problems, preprocessing with EFS, some original types of EAs and its application of EFSs, and the management of the intrinsic data characteristics to better adapt to the context of each individual problem.

In summary, the objectives of this work were to highlight the good properties of EFSs to be applied as "multipurpose" tools, and to advise researchers on topics in the near future, and which is most the appropriate methodology to follow when releasing new proposals on the topic.

Acknowledgments

This work was partially supported by the Spanish Ministry of Science and Technology under projects TIN2011-28488, TIN-2012-33856, and the Andalusian Research Plans P12-TIC-2958, P11-TIC-7765 and P10-TIC-6858.

References

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

[2] G. Abu-Lebdeh, R.F. Benekohal, Convergence variability and population sizing in micro-genetic algorithms, Comput.-Aided Civ. Infrastruct. Eng. 14 (5) (1999) 321–334.

[3] R. Agrawal, R. Srikant, Fast algorithms for mining association rules, in: J.B. Bocca, M. Jarke, C. Zaniolo (Eds.), Proceedings of 20th International Conference on Very Large Data Bases, VLDB, Morgan Kaufman, 1994, pp. 487–499.

[4] E. Alba, J.M. Troya, Improving flexibility and efficiency by adding parallelism to genetic algorithms, Stat. Comput. 12 (2) (2002) 91–114.

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

[6] R. Alcala, M.J. Gacto, F. Herrera, A fast and scalable multiobjective genetic fuzzy system for linguistic fuzzy modeling in high-dimensional regression problems, IEEE Trans. Fuzzy Syst. 19 (4) (2011) 666–681.

[7] J. Alcala-Fdez, R. Alcala, M.J. Gacto, F. Herrera, Learning the membership function contexts for mining fuzzy association rules by using genetic algorithms, Fuzzy Sets Syst. 160 (7) (2009) 905–921.

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

[9] J. Alcala-Fdez, A. Fernandez, J. Luengo, J. Derrac, S. Garcia, L. Sanchez, F. Herrera, KEEL data-mining software tool: data set repository, integration of algorithms and experimental analysis framework, J. Multi-Valued Logic Soft Comput. 17 (2–3) (2011) 255–287.

[10] J. Alcala-Fdez, F. Herrera, F.A. Marquez, A. Peregrin, Increasing fuzzy rules cooperation based on evolutionary adaptive inference systems, Int. J. Intell. Syst. 22 (9) (2007) 1035–1064.

[11] J. Alcala;-Fdez, L. Sanchez, S. Garcia, M.J. del Jesus, S. Ventura, J.M. Garrell, J. Otero, C. Romero, J. Bacardit, V.M. Rivas, J.C. Fernandez, F. Herrera, KEEL: a software tool to assess evolutionary algorithms for data mining problems, Soft Comput. 13 (2009) 307–318.

[12] S. Alshomrani, A. Bawakid, S.O. Shim, A. Fernandez, F. Herrera, A proposal for evolutionary fuzzy systems using feature weighting: dealing with overlapping in imbalanced datasets, Knowl.-Based Syst. 73 (2015) 1–17.

[13] A. Alvarez-Alvarez, G. Trivino, O. Cordon, Human gait modeling using a genetic fuzzy finite state machine, IEEE Trans. Fuzzy Syst. 20 (2) (2012) 205–223.

[14] J. Amores, Multiple instance classification: review, taxonomy and comparative study, Artif. Intell. 201 (2013) 81–105.

[15] M. Antonelli, P. Ducange, F. Marcelloni, Genetic training instance selection in multiobjective evolutionary fuzzy systems: a coevolutionary approach, IEEE Trans. Fuzzy Syst. 20 (2) (2012) 276–290.

[16] A. Azhar Ramli, J. Watada, W. Pedrycz, A combination of genetic algorithm-based fuzzy c-means with a convex hull-based regression for real-time fuzzy switching regression analysis: application to industrial intelligent data analysis, IEEE Trans. Electr. Electron. Eng. 9 (1) (2014) 71–82.

[17] J.L. Aznarte, J. Alcala-Fdez, A. Arauzo, J.M. Benitez, Financial time series forecasting with a bio-inspired fuzzy model, Expert Syst. Appl. 39 (16) (2013) 12302–12309.

[18] J. Bacardit, E.K. Burke, N. Krasnogor, Improving the scalability of rule-based evolutionary learning, Memetic Comput. 1 (1) (2009) 55–67.

[19] G. Beliakov, S. James, Citation-based journal ranks: the use of fuzzy measures, Fuzzy Sets Syst. 167 (1) (2011) 101–119.

[20] A.D. Benitez, J. Casillas, Multi-objective genetic learning of serial hierarchical fuzzy systems for large-scale problems, Soft Comput. 17 (1) (2013) 165–194.

[21] H.R. Berenji, P.S. Khedkar, Learning and tuning fuzzy logic controllers through reinforcements, IEEE Trans. Neural Netw. 3 (5) (1992) 724–740.

[22] J.Q. Candela, M. Sugiyama, A. Schwaighofer, N.D. Lawrence, Dataset Shift in Machine Learning, The MIT Press, 2009.

[23] J.S. Cardoso, R. Sousa, Measuring the performance of ordinal classification, Int. J. Pattern Recognit. Artif. Intell. 25 (8) (2011) 1173–1195.

[24] C.J. Carmona, C. Chrysostomou, H. Seker, M.J. del Jesus, Fuzzy rules for describing subgroups from influenza a virus using a multi-objective evolutionary algorithm, Appl. Soft Comput. 13 (2013) 3439–3448.

[25] C.J. Carmona, P. Gonzalez, M.J. del Jesus, F. Herrera, NMEEF-SD: non-dominated multi-objective evolutionary algorithm for extracting fuzzy rules in subgroup discovery, IEEE Trans. Fuzzy Syst. 18 (5) (2010) 958–970.

[26] C.J. Carmona, P. Gonzalez, M.J. del Jesus, M. Navio-Acosta, L. Jimenez-Trevino, Evolutionary fuzzy rule extraction for subgroup discovery in a psychiatric emergency department, Soft Comput. 15 (12) (2011) 2435–2448.

[27] C.J. Carmona, P. Gonzalez, B. Garcia-Domingo, M.J. del Jesus, J. Aguilera, MEFES: an evolutionary proposal for the detection of exceptions in subgroup discovery. An application to concentrating photovoltaic technology, Knowl.-Based Syst. 54 (2013) 73–85.

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

[29] O. Castillo, P. Melin, Optimization of type-2 fuzzy systems based on bio-inspired methods: a concise review, Inf. Sci. 205 (2012) 1–19.

[30] O. Castillo, P. Melin, A.A. Garza, O. Montiel, R. Sepulveda, Optimization of interval type-2 fuzzy logic controllers using evolutionary algorithms, Soft Comput. 15 (6) (2011) 1145–1160.

[31] N.V. Chawla, K.W. Bowyer, L.O. Hall, W.P. Kegelmeyer, SMOTE: synthetic minority over-sampling technique, J. Artif. Intell. Res. 16 (2002) 321–357.

[32] C.A. Coello-Coello, G. Lamont, D. van Veldhuizen, Evolutionary Algorithms for Solving Multi-objective Problems, second ed., Genetic and Evolutionary Computation, Springer, Berlin, Heidelberg, 2007.

[33] C.A. Coello-Coello, G.T. Pulido, A micro-genetic algorithm for multiobjective optimization, in: E. Zitzler, K. Deb, L. Thiele, C.A. Coello-Coello, D. Corne (Eds.), Proceedings of the First International Conference on Evolutionary Multi-Criterion Optimization (EMO'2001), Lecture Notes in Computer Science, vol. 1993, Springer, 2001, pp. 126–140.

[34] W.E. Combs, J.E. Andrews, Combinatorial rule explosion eliminated by a fuzzy rule configuration, IEEE Trans. Fuzzy Syst. 6 (1) (1998) 1–11.

[35] R.L.F. Cordeiro, C. Faloutsos, C. Traina Jr., Data Mining in Large Sets of Complex Data, Springer Briefs in Computer Science, Springer, 2013.

[36] 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. 52 (6) (2011) 894–913.

[37] O. Cordon, M. del Jesus, F. Herrera, M. Lozano, MOGUL: a methodology to obtain genetic fuzzy rule-based systems under the iterative rule learning approach, Int. J. Intell. Syst. 14 (1999) 123–1153.

[38] O. Cordon, F. Gomide, F. Herrera, F. Hoffmann, L. Magdalena, Ten years of genetic fuzzy systems: current framework and new trends, Fuzzy Sets Syst. 141 (2004) 5–31.

[39] O. Cordon, F. Herrera, F. Hoffmann, L. Magdalena, Genetic Fuzzy Systems. Evolutionary Tuning and Learning of Fuzzy Knowledge Bases, World Scientific, Singapore, Republic of Singapore, 2001.

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

[41] M. Cruz-Ramirez, C. Hervas-Martinez, J. Sanchez-Monedero, P.A. Gutierrez, Metrics to guide a multi-objective evolutionary algorithm for ordinal classification, Neurocomputing 135 (2014) 21–31.

[42] A. de Haro-Garcia, N. Garcia-Pedrajas, A divide-and-conquer recursive approach for scaling up instance selection algorithms, Data Min. Knowl. Disc. 18 (3) (2009) 392–418.

[43] K. Deb, H. Jain, An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, Part I: solving

problems with box constraints, IEEE Trans. Evol. Comput. 18 (4) (2014) 577–601.
[44] M.J. del Jesus, P. Gonzalez, F. Herrera, M. Mesonero, Evolutionary fuzzy rule induction process for subgroup discovery: a case study in marketing, IEEE Trans. Fuzzy Syst. 15 (4) (2007) 578–592.
[45] J. Demšar, Statistical comparisons of classifiers over multiple data sets, J. Mach. Learn. Res. 7 (2006) 1–30.
[46] T. Dietterich, R. Lathrop, T. Lozano-Perez, Solving the multiple instance problem with axis-parallel rectangles, Artif. Intell. 89 (1–2) (1997) 31–71.
[47] W. Dong, Z. Huang, L. Ji, H. Duan, A genetic fuzzy system for unstable angina risk assessment, BMC Med. Inform. Decis. Mak. 14 (1) (2014).
[48] D. Dubois, H. Prade, T. Sudkamp, On the representation, measurement, and discovery of fuzzy associations, IEEE Trans. Fuzzy Syst. 13 (2005) 250–262.
[49] P. Ducange, B. Lazzerini, F. Marcelloni, Multi-objective genetic fuzzy classifiers for imbalanced and cost-sensitive datasets, Soft Comput. 14 (7) (2010) 713–728.
[50] A.E. Eiben, J.E. Smith, Introduction to Evolutionary Computation, Springer-Verlag, Berlin, Germany, 2003.
[51] S. Elhag, A. Fernandez, A. Bawakid, S. Alshomrani, F. Herrera, On the combination of genetic fuzzy systems and pairwise learning for improving detection rates on intrusion detection systems, Expert Syst. Appl. 42 (1) (2015) 193–202.
[52] M. Fahim, I. Fatima, S. Lee, Y.T. Park, EFM: evolutionary fuzzy model for dynamic activities recognition using a smartphone accelerometer, Appl. Intell. 39 (3) (2013) 475–488.
[53] M. Fazzolari, R. Alcala, Y. Nojima, H. Ishibuchi, F. Herrera, A review of the application of multi-objective evolutionary systems: current status and further directions, IEEE Trans. Fuzzy Syst. 21 (1) (2013) 45–65.
[54] M. Fazzolari, B. Giglio, R. Alcala, F. Marcelloni, F. Herrera, A study on the application of instance selection techniques in genetic fuzzy rule-based classification systems: accuracy-complexity trade-off, Knowl.-Based Syst. 54 (2013) 32–41.
[55] A. Fernandez, M.J. del Jesus, F. Herrera, Hierarchical fuzzy rule based classification systems with genetic rule selection for imbalanced data-sets, Int. J. Approx. Reason. 50 (2009) 561–577.
[56] A. Fernandez, M.J. del Jesus, F. Herrera, On the influence of an adaptive inference system in fuzzy rule based classification systems for imbalanced data-sets, Expert Syst. Appl. 36 (6) (2009) 9805–9812.
[57] A. Fernandez, M.J. del Jesus, F. Herrera, On the 2-tuples based genetic tuning performance for fuzzy rule based classification systems in imbalanced data-sets, Inf. Sci. 180 (8) (2010) 1268–1291.
[58] A. Fernandez, S. del Rio, V. Lopez, A. Bawakid, M.J. del Jesus, J.M. Benitez, F. Herrera, Big data with cloud computing: an insight on the computing environment, MapReduce and programming frameworks, Wiley Interdiscipl. Rev.: Data Min. Knowl. Discov. 4 (5) (2014) 380–409.
[59] A. Fernandez, S. Garcia, M.J. del Jesus, F. Herrera, A study of the behaviour of linguistic fuzzy rule based classification systems in the framework of imbalanced data-sets, Fuzzy Sets Syst. 159 (18) (2008) 2378–2398.
[60] B. Frenay, M. Verleysen, Classification in the presence of label noise: a survey, IEEE Trans. Neural Netw. Learn. Syst. 25 (5) (2014) 845–869.
[61] M.J. Gacto, R. Alcala, F. Herrera, Adaptation and application of multi-objective evolutionary algorithms for rule reduction and parameter tuning of fuzzy rule-based systems, Soft Comput. 13 (5) (2009) 419–436.
[62] M.J. Gacto, R. Alcala, F. Herrera, Interpretability of linguistic fuzzy rule-based systems: an overview of interpretability measures, Inf. Sci. 181 (20) (2011) 4340–4360.
[63] M.J. Gacto, M. Galende, R. Alcala, F. Herrera, METSK-HDe: a multiobjective evolutionary algorithm to learn accurate TSK-fuzzy systems in high-dimensional and large-scale regression problems, Inf. Sci. 276 (2014) 63–79.
[64] J. Gama, M. Gaber (Eds.), Learning from Data Streams – Processing Techniques in Sensor Networks, Springer, 2007.
[65] D. Garcia, A. Gonzalez, R. Perez, Overview of the SLAVE learning algorithm: a review of its evolution and prospects, Int. J. Comput. Intell. Syst. 7 (6) (2014) 1194–1221.
[66] S. Garcia, A. Fernandez, J. Luengo, F. Herrera, A study of statistical techniques and performance measures for genetics-based machine learning: accuracy and interpretability, Soft Comput. 13 (10) (2009) 959–977.
[67] S. Garcia, A. Fernandez, J. Luengo, F. Herrera, Advanced nonparametric tests for multiple comparisons in the design of experiments in computational intelligence and data mining: experimental analysis of power, Inf. Sci. 180 (2010) 2044–2064.
[68] S. Garcia, F. Herrera, An extension on "statistical comparisons of classifiers over multiple data sets" for all pairwise comparisons, J. Mach. Learn. Res. 9 (2008) 2677–2694.
[69] S. Garcia, J. Luengo, F. Herrera, Data Preprocessing in Data Mining, Intelligent Systems Reference Library, first ed., vol. 72, Springer, 2014.
[70] V. Garcia, R.A. Mollineda, J.S. Sanchez, On the k-NN performance in a challenging scenario of imbalance and overlapping, Pattern Anal. Appl. 11 (3–4) (2008) 269–280.
[71] N. Garcia-Pedrajas, A. de Haro-Garcia, Scaling up data mining algorithms: review and taxonomy, Progr. Artif. Intell. 1 (1) (2012) 71–87.
[72] N. Garcia-Pedrajas, A. Haro, J. Perez, A scalable approach to simultaneous evolutionary instance and feature selection, Inf. Sci. 228 (2013) 150–174.
[73] S. Guillaume, Designing fuzzy inference systems from data: an interpretability-oriented review, IEEE Trans. Fuzzy Syst. 9 (3) (2001) 426–443.
[74] J. Han, M. Kamber, J. Pei, Data Mining: Concepts and Techniques, third ed., Morgan Kaufman, San Mateo, CA, USA, 2011.
[75] F. Herrera, Genetic fuzzy systems: taxonomy, current research trends and prospects, Evol. Intel. 1 (1) (2008) 27–46.
[76] F. Herrera, C. Carmona, P. Gonzalez, M.J. del Jesus, An overview on subgroup discovery: foundations and applications, Knowl. Inf. Syst. (2010) 1–31.
[77] A. Homaifar, E. McCormick, Simultaneous design of membership functions and rule sets for fuzzy controllers using genetic algorithms, IEEE Trans. Fuzzy Syst. 3 (2) (1995) 129–139.
[78] T.P. Hong, C.H. Chen, Y.C. Lee, Y.L. Wu, Genetic-fuzzy data mining with divide-and-conquer strategy, IEEE Trans. Evol. Comput. 12 (2) (2008) 252–265.
[79] Q. Hu, W. Pan, L. Zhang, D. Zhang, Y. Song, M. Guo, D. Yu, Feature selection for monotonic classification, IEEE Trans. Fuzzy Syst. 20 (1) (2012) 69–81.
[80] O. Ibañez, O. Cordon, S. Damas, A cooperative coevolutionary approach dealing with the skull-face overlay uncertainty in forensic identification by craniofacial superimposition, Soft Comput. 16 (5) (2012) 797–808.
[81] H. Ishibuchi, T. Murata, I. Turksen, Single-objective and two-objective genetic algorithms for selecting linguistic rules for pattern classification problems, Fuzzy Sets Syst. 8 (2) (1997) 135–150.
[82] H. Ishibuchi, T. Nakashima, M. Nii, Classification and Modeling with Linguistic Information Granules: Advanced Approaches to Linguistic Data Mining, Springer-Verlag, Berlin, Germany, 2004.
[83] H. Ishibuchi, K. Nozaki, N. Yamamoto, H. Tanaka, Selection of fuzzy IF-THEN rules for classification problems using genetic algorithms, IEEE Trans. Fuzzy Syst. 3 (3) (1995) 260–270.
[84] H. Ishibuchi, T. Yamamoto, T. Nakashima, Hybridization of fuzzy GBML approaches for pattern classification problems, IEEE Trans. Syst. Man, Cybern. – Part B 35 (2) (2005) 359–365.
[85] H. Jain, K. Deb, An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, Part II: handling constraints and extending to an adaptive approach, IEEE Trans. Evol. Comput. 18 (4) (2014) 602–622.
[86] J.Y. Jiang, S.C. Tsai, S.J. Lee, FSKNN: multi-label text categorization based on fuzzy similarity and k nearest neighbors, Expert Syst. Appl. 39 (3) (2012) 2813–2821.
[87] Y. Jin, Fuzzy modeling of high-dimensional systems: complexity reduction and interpretability improvement, IEEE Trans. Fuzzy Syst. 8 (2) (2000) 212–221.
[88] C.-F. Juang, C.-T. Lin, An online self-constructing neural fuzzy inference network and its applications, IEEE Trans. Fuzzy Syst. 6 (1) (1998) 12–32.
[89] H. Karau, A. Konwinski, P. Wendell, M. Zaharia, Learning Spark: Lightning-Fast Big Data Analytics, first ed., O'Reilly Media, 2014.
[90] N.N. Karnik, J.M. Mendel, Q. Liang, Type-2 fuzzy logic systems, IEEE Trans. Fuzzy Syst. 7 (6) (1999) 643–658.
[91] D. Kim, Y. Choi, S.-Y. Lee, An accurate COG defuzzifier design using lamarckian co-adaptation of learning and evolution, Fuzzy Sets Syst. 130 (2) (2002) 207–225.
[92] W. Kotlowski, R. Slowinski, On nonparametric ordinal classification with monotonicity constraints, IEEE Trans. Knowl. Data Eng. 25 (11) (2013) 2576–2589.
[93] C. Lam, Hadoop in Action, first ed., Manning, 2011.
[94] N. Lavrac, B. Cestnik, D. Gamberger, P.A. Flach, Decision support through subgroup discovery: three case studies and the lessons learned, Mach. Learn. 57 (1–2) (2004) 115–143.
[95] W.-P. Lee, Y.-T. Hsiao, W.-C. Hwang, Designing a parallel evolutionary algorithm for inferring gene networks on the cloud computing environment, BMC Syst. Biol. 8 (1) (2014) 1–5.
[96] C.-T. Lin, C.S.G. Lee, Neural-network-based fuzzy logic control and decision system, IEEE Trans. Comput. 40 (12) (1991) 1320–1336.
[97] C.F. Liu, C.Y. Yeh, S.J. Lee, Application of type-2 neuro-fuzzy modeling in stock price prediction, Appl. Soft Comput. J. 12 (4) (2012) 1348–1358.
[98] V. Lopez, S. del Rio, J.M. Benitez, F. Herrera, On the use of MapReduce to build linguistic fuzzy rule based classification systems for big data, in: IEEE World Congress on Computational Intelligence (WCCI 2014). IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2014), Beijing (China), 2014, pp. 1905–1912.
[99] V. Lopez, S. del Rio, J.M. Benitez, F. Herrera, Cost-sensitive linguistic fuzzy rule based classification systems under the MapReduce framework for imbalanced big data, Fuzzy Sets Syst. 258 (2015) 5–38.
[100] V. Lopez, A. Fernandez, M.J. del Jesus, F. Herrera, A hierarchical genetic fuzzy system based on genetic programming for addressing classification with highly imbalanced and borderline data-sets, Knowl. Based Syst. 38 (2013) 85–104.
[101] V. Lopez, A. Fernandez, S. Garcia, V. Palade, F. Herrera, An insight into classification with imbalanced data: empirical results and current trends on using data intrinsic characteristics, Inf. Sci. 250 (20) (2013) 113–141.
[102] V. Lopez, A. Fernandez, F. Herrera, Addressing covariate shift for genetic fuzzy systems classifiers: a case of study with FARC-HD for imbalanced datasets, in: Proceedings of the 2013 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2013), 2013, pp. 1–8.
[103] V. Lopez, A. Fernandez, F. Herrera, On the importance of the validation technique for classification with imbalanced datasets: addressing covariate shift when data is skewed, Inf. Sci. 257 (2014) 1–13.
[104] J. Luengo, A. Fernandez, S. Garcia, F. Herrera, Addressing data complexity for imbalanced data sets: analysis of SMOTE–based oversampling and evolutionary undersampling, Soft Comput. 15 (10) (2011) 1909–1936.

[105] S.S. Madaeni, A.R. Kurdian, Fuzzy modeling and hybrid genetic algorithm optimization of virus removal from water using microfiltration membrane, Chem. Eng. Res. Des. 89 (4) (2011) 456–470.
[106] A. Mahnot, M. Popescu, FUMIL-fuzzy multiple instance learning for early illness recognition in older adults, in: 2012 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), 2012, pp. 1–5.
[107] A.A. Marquez, F.A. Marquez, A.M. Roldan, A. Peregrin, An efficient adaptive fuzzy inference system for complex and high dimensional regression problems in linguistic fuzzy modelling, Knowl.-Based Syst. 54 (2013) 42–52.
[108] F. Marquez, A. Peregrín, F. Herrera, Cooperative evolutionary learning of linguistic fuzzy rules and parametric aggregation connectors for Mamdani fuzzy systems, IEEE Trans. Fuzzy Syst. 15 (6) (2008) 1162–1178.
[109] M. Minelli, M. Chambers, A. Dhiraj, Big Data, first ed., Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses, Wiley, 2013.
[110] J.G. Moreno-Torres, T. Raeder, R. Alaiz-Rodriguez, N.V. Chawla, F. Herrera, A unifying view on dataset shift in classification, Pattern Recogn. 45 (1) (2012) 521–530.
[111] J.G. Moreno-Torres, J.A. Saez, F. Herrera, Study on the impact of partition-induced dataset shift on k-fold cross-validation, IEEE Trans. Neural Netw. Learn. Syst. 23 (8) (2012) 1304–1313.
[112] A. Mukhopadhyay, U. Maulik, S. Bandyopadhyay, C.A. Coello-Coello, A survey of multiobjective evolutionary algorithms for data mining: Part I, IEEE Trans. Evol. Comput. 18 (1) (2014) 4–19.
[113] S.K. Mylonas, D.G. Stavrakoudis, J.B. Theocharis, Genesis: a GA-based fuzzy segmentation algorithm for remote sensing images, Knowl.-Based Syst. 54 (2013) 86–102.
[114] D. Nauck, R. Kruse, A neuro-fuzzy method to learn fuzzy classification rules from data, Fuzzy Sets Syst. 89 (3) (1997) 277–288.
[115] M.T. Nouei, A.V. Kamyad, M. Sarzaeem, S. Ghazalbash, Developing a genetic fuzzy system for risk assessment of mortality after cardiac surgery, J. Med. Syst. 38 (102) (2014) 1–9.
[116] A. Orriols-Puig, J. Casillas, Fuzzy knowledge representation study for incremental learning in data streams and classification problems, Soft Comput. 15 (12) (2011) 2389–2414.
[117] A. Orriols-Puig, F.J. Martinez-Lopez, J. Casillas, A soft-computing-based method for the automatic discovery of fuzzy rules in databases: uses for academic research and management support in marketing, J. Bus. Res. 66 (9) (2013) 1332–1337.
[118] A. Orriols-Puig, F.J. Martinez-Lopez, J. Casillas, N. Lee, Unsupervised {KDD} to creatively support managers' decision making with fuzzy association rules: a distribution channel application, Ind. Mark. Manage. 42 (4) (2013) 532–543.
[119] A. Palacios, L. Sanchez, I. Couso, Diagnosis of dyslexia with low quality data with genetic fuzzy systems, Int. J. Approx. Reason. 51 (8) (2010) 993–1009.
[120] A. Palacios, L. Sanchez, I. Couso, Future performance modeling in athletism with low quality data-based genetic fuzzy systems, J. Multiple-Valued Logic Soft Comput. 17 (2–3) (2011) 207–228.
[121] A. Palacios, L. Sanchez, I. Couso, Linguistic cost-sensitive learning of genetic fuzzy classifiers for imprecise data, Int. J. Approx. Reason. 52 (6) (2011) 841–862.
[122] A. Palacios, K. Trawinski, O. Cordon, L. Sanchez, Cost-sensitive learning of fuzzy rules for imbalanced classification problems using FURIA, Int. J. Uncertain. Fuzz. Knowl.-Based Syst. 20 (2014) 1–20.
[123] W. Pedrycz, F. Gomide, Fuzzy Systems Engineering: Toward Human-Centric Computing, first ed., Wiley, 2007.
[124] E. Perez, M. Posada, F. Herrera, Analysis of new niching genetic algorithms for finding multiple solutions in the job shop scheduling, J. Intell. Manuf. 23 (3) (2012) 341–356.
[125] F.J. Provost, V. Kolluri, A survey of methods for scaling up inductive algorithms, Data Min. Knowl. Disc. 3 (2) (1999) 131–169.
[126] J.R., Neuro-fuzzy Modeling: Architectures, Analyses and Applications, Ph.D. Thesis, University of California, Berkeley, 1992.
[127] R. Sambuc, Function ϕ-flous, application a l'aide au diagnostic en pathologie thyroidienne, Ph.D. thesis, University of Marseille, 1975.
[128] S.J. Raudys, A.K. Jain, Small sample size effects in statistical pattern recognition: recommendations for practitioners, IEEE Trans. Pattern Anal. Mach. Intell. 13 (3) (1991) 252–264.
[129] I. Robles, R. Alcala, J.M. Benitez, F. Herrera, Evolutionary parallel and gradually distributed lateral tuning of fuzzy rule-based systems, Evol. Intel. 2 (1–2) (2009) 5–19.
[130] J.A. Saez, J. Luengo, F. Herrera, Fuzzy rule based classification systems versus crisp robust learners trained in presence of class noise's effects: a case of study, in: 11th International Conference on Intelligent Systems Design and Applications (ISDA), 2011, pp. 1229–1234.
[131] J.A. Saez, J. Luengo, J. Stefanowski, F. Herrera, SMOTE-IPF: addressing the noisy and borderline examples problem in imbalanced classification by a re-sampling method with filtering, Inf. Sci. 291 (2015) 184–203.
[132] D. Sanchez, C. Cornelis, R.E. Bello, F. Herrera, Multi-instance learning wrapper based on the Rocchio classifier for web index recommendation, Knowl.-Based Syst. 59 (2014) 173–181.
[133] J.S. Sanchez, R.A. Mollineda, J.M. Sotoca, An analysis of how training data complexity affects the nearest neighbor classifiers, Pattern Anal. Appl. 10 (3) (2007) 189–201.
[134] L. Sanchez, J. Casillas, O. Cordon, M.J. del Jesus, Some relationships between fuzzy and random set-based classifiers and models, Int. J. Approx. Reason. 29 (2) (2002) 175–213.
[135] L. Sanchez, I. Couso, J. Casillas, Genetic learning of fuzzy rules based on low quality data, Fuzzy Sets Syst. 160 (17) (2009) 2524–2552.
[136] L. Sanchez, I. Couso, A. Palacios, J. Palacios, A methodology for exploiting the tolerance for imprecision in genetic fuzzy system and its application to characterization of rotor blade learning edge materials, IEEE Trans. Evol. Comput. 37 (1–2) (2013) 76–91.
[137] J. Sanz, D. Bernardo, F. Herrera, H. Bustince, 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. (2014) 1–14 (in press). http://dx.doi.org/10.1109/TFUZZ.2014.2336263.
[138] J. Sanz, A. Fernandez, H. Bustince, F. Herrera, A genetic tuning to improve the performance of fuzzy rule-based classification systems with interval-valued fuzzy sets: degree of ignorance and lateral position, Int. J. Approx. Reason. 52 (6) (2011) 751–766.
[139] J.A. Sanz, A. Fernandez, H. Bustince, F. Herrera, Improving the performance of fuzzy rule-based classification systems with interval-valued fuzzy sets and genetic amplitude tuning, Inf. Sci. 180 (19) (2010) 3674–3685.
[140] J.A. Sanz, A. Fernandez, H. Bustince, F. Herrera, IVTURS: a linguistic fuzzy rule-based classification system based on a new interval-valued fuzzy reasoning method with tuning and rule selection, IEEE Trans. Fuzzy Syst. 21 (3) (2013) 399–411.
[141] L. Septem Riza, C. Bergmeir, F. Herrera, J. Benitez, Learning from data using the r package "frbs", in: 2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), 2014, pp. 2149–2155.
[142] F. Serdio, E. Lughofer, K. Pichler, T. Buchegger, H. Efendic, Residual-based fault detection using soft computing techniques for condition monitoring at rolling mills, Inf. Sci. 259 (2014) 304–320.
[143] L. Sheng, X. Ma, A novel GA-fuzzy classification method for audio signals, J. Inform. Comput. Sci. 9 (3) (2012) 595–610.
[144] V. Soler, M. Prim, Extracting a fuzzy system by using genetic algorithms for imbalanced datasets classification: application on down's syndrome detection, in: D.A. Zighed, S. Tsumoto, Z.W. Ras, H. Hacid (Eds.), Mining Complex Data, Studies in Computational Intelligence, vol. 165, Springer, 2009, pp. 23–39.
[145] D.G. Stavrakoudis, J.B. Theocharis, G.C. Zalidis, A boosted genetic fuzzy classifier for land cover classification of remote sensing imagery, ISPRS J. Photogramm. Remote Sens. 66 (4) (2011) 529–544.
[146] R. Sumathi, E. Kirubakaran, R. Krishnamoorthy, Multi class multi label based fuzzy associative classifier with genetic rule selection for coronary heart disease risk level prediction, Int. Rev. Comput. Softw. 9 (3) (2014) 533–540.
[147] T. Takagi, M. Sugeno, Fuzzy identification of systems and its applications to modeling and control, IEEE Trans. Syst. Man Cybern. 15 (1) (1985) 116–132.
[148] S. Tano, T. Oyama, T. Arnould, Deep combination of fuzzy inference and neural network in fuzzy inference software – {FINEST}, Fuzzy Sets Syst. 82 (2) (1996) 151–160.
[149] P. Thrift, Fuzzy logic synthesis with genetic algorithms, in: Proceedings of the 4th International Conference on Genetic Algorithms (ICGA'91), 1991, pp. 509–513.
[150] B. Trawinski, Evolutionary fuzzy system ensemble approach to model real estate market based on data stream exploration, J. Univ. Comput. Sci. 19 (4) (2013) 539–562.
[151] M.E. Turan, M.A. Yurdusev, Predicting monthly river flows by genetic fuzzy systems, Water Resour. Manage. 28 (13) (2014) 4685–4697.
[152] T.A. Victoire, M. Sakthivel, A refined differential evolution algorithm based fuzzy classifier for intrusion detection, Eur. J. Sci. Res. 65 (2) (2011) 246–259.
[153] P. Villar, A. Fernandez, R.A. Carrasco, F. Herrera, Feature selection and granularity learning in genetic fuzzy rule-based classification systems for highly imbalanced data-sets, Int. J. Uncertain. Fuzz. Knowl. Based Syst. 20 (3) (2012) 369–397.
[154] C. Wagner, H. Hagras, A genetic algorithm based architecture for evolving type-2 fuzzy logic controllers for real world autonomous mobile robots, in: FUZZ-IEEE, IEEE, 2007, pp. 1–6.
[155] M. Wasikowski, X.-W. Chen, Combating the small sample class imbalance problem using feature selection, IEEE Trans. Knowl. Data Eng. 22 (10) (2010) 1388–1400.
[156] G.M. Weiss, Mining with rarity: a unifying framework, SIGKDD Explor. 6 (1) (2004) 7–19.
[157] G.M. Weiss, The impact of small disjuncts on classifier learning, in: R. Stahlbock, S.F. Crone, S. Lessmann (Eds.), Data Mining, Annals of Information Systems, vol. 8, Springer, 2010, pp. 193–226.
[158] G.M. Weiss, F.J. Provost, Learning when training data are costly: the effect of class distribution on tree induction, J. Artif. Intell. Res. 19 (2003) 315–354.
[159] L.A. Zadeh, Fuzzy sets, Inf. Control 8 (1965) 338–353.
[160] C. Zhang, S. Zhang, in: Association Rule Mining, Models and Algorithms, Lecture Notes in Computer Science, vol. 2307, Springer, 2002.
[161] M.-L. Zhang, Z.-H. Zhou, A review on multi-label learning algorithms, IEEE Trans. Knowl. Data Eng. 26 (8) (2014) 1819–1837.
[162] Z.-H. Zhou, M.-L. Zhang, S.-J. Huang, Y.-F. Li, Multi-instance multi-label learning, Artif. Intell. 176 (1) (2012) 2291–2320.
[163] X. Zhu, X. Wu, Class noise vs. attribute noise: a quantitative study, Artif. Intell. Rev. 22 (3) (2004) 177–210.
[164] P.C. Zikopoulos, C. Eaton, D. deRoos, T. Deutsch, G. Lapis, Understanding Big Data – Analytics for Enterprise Class Hadoop and Streaming Data, first ed., McGraw-Hill Osborne Media, 2011.