Elsevier Editorial System(tm) for Applied Soft Computing
Manuscript Draft

Manuscript Number: ASOC-D-19-04331R1

Title: A wrapper methodology to learn interval-valued fuzzy rule-based classification systems

Article Type: Full Length Article

Keywords: Classification problems; Evolutionary fuzzy systems; Fuzzy rule-based classification systems; Interval-valued fuzzy sets; Interval type-2 fuzzy sets

Corresponding Author: Dr. Jose Antonio Sanz,

Corresponding Author's Institution: Universidad Publica de Navarra

First Author: Jose Antonio Sanz

Order of Authors: Jose Antonio Sanz; Humberto Bustince

Abstract: Learning an interval-valued fuzzy rule-based classification system is a challenge as its success directly depends on the interval-valued fuzzy partition used. In fact, the learning of an interval-valued fuzzy system usually starts by creating a partition composed of numerical fuzzy sets, which are used to build an initial fuzzy classifier. Then, it is augmented with interval-valued fuzzy sets whose shape is subsequently optimized to improve the system's performance. However, as in this methodology the fuzzy rules are learned using numerical fuzzy sets, the benefits of the interval-valued fuzzy sets may not be fully exploited. In this paper we define a new learning methodology that avoids building the initial fuzzy classifier but directly learns interval-valued fuzzy rules. To do so, we define a wrapper methodology to learn the interval-valued fuzzy partitions such that they lead to an interval-valued fuzzy rule-based classification system as accurate as possible. Moreover, our new method allows one to represent each membership function using the most proper type of fuzzy set for the sake of modeling the uncertainty in the best possible manner. Consequently, the antecedents of the rules can be formed of only numerical fuzzy sets, only interval-valued fuzzy sets or a mixture of both. The quality of the proposal is compared versus four state-of-the-art fuzzy classifiers like FARC-HD, IVTURS, FURIA and FARC-HD using an inference based on a generalization of the Choquet integral. We also compare our new approach besides its numerical fuzzy counterpart to clearly show the benefits of the usage of interval-valued fuzzy sets. Specifically, the average accuracy rate of our new method is 81.17\%, which is at least 0.66\% better than the remainder state-of-the-art fuzzy classifiers.

***Declaration of Interest Statement**

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Declarations of interest: none

Dear Prof. M. Köppen,

Please find enclosed a revised version of the manuscript entitled "*A wrapper methodology to learn interval-valued fuzzy rule-based classification systems*" and its authors are J. Sanz and H. Bustince. The content of the paper is the authors' original work and has not been published nor has it been submitted simultaneously elsewhere. Both authors have checked the manuscript and have agreed to the submission.

Sincerely yours,

Jose Antonio

Dr. Jose Antonio Sanz

***Detailed Response to Reviewers**

Applied Soft Computing

Re: ASOC-D-19-04331

A wrapper methodology to learn interval-valued fuzzy rule-based classification systems

J. Sanz, H. Bustince.

Authors' Response letter

We wish to sincerely thank the reviewers for their comments, since they have allowed us to improve the overall quality of the paper.

In the following we first describe the main changes introduced in the revised paper and then we provide detailed responses to the referees.

Main Changes

- We have rewritten the highlights to fulfil the journal's requirements.
- We have stressed the accuracy of IVFARC and its improvement versus the remainder methods in the abstract.
- We have clarified the steps of the new proposal both in the introduction and in Section 3, as well as some features of the evolutionary process.
- We have included a recent fuzzy classifier, which is related to our new method, in the comparison in order to support the quality of IVFARC. Specifically, we have added a novel fuzzy classifier that uses FARC-HD to construct the fuzzy rules but it applies a generalization of the Choquet integral to fuse the information of the fired rules in the inference process (reference [25] of the new version of the paper). Consequently, we have modified both the text and the tables, to introduce its set-up (Section 4), its results and to analyse it (Section 5).
- We have divided the previous section 5.1 of the experimental study into four sections looking for a better organization of the analysis. Furthermore, we have included a new subsection to analyse the run-time of the classifiers (new section 5.2.2).
- We have elaborated more on the future research lines pointed out in the Conclusions section (section 6).
- We have updated the references by including new recent related papers to our new approach.

REVIEWER 1

RV: Reviewer
AA: Author Answers

RV: Authors are suggested to address the following comments. 1. Abstract, please state the performance of proposed algorithm numerically. Also, state the percentage improvement by proposed algorithm compared with existing works.
AA: Thanks for this comment because in this way the performance of the new method and its improvement is stated from the beginning. We have modified the final sentences of the abstract in order to include this information.

RV: 2. Highlights, please address: (i) Follow journal's guideline, "maximum 85 characters, including spaces, per bullet point". (ii) State the key benefits of proposed algorithm. (iii) State the percentage improvement by proposed algorithm compared with existing works.
AA: Thanks for the comment. We have modified the highlights so that they fulfil the journal’s requirements and we have highlighted the minimum percentage of improvement of our method.

RV: 3. Section 1, authors provide a good background and history of the topic. However, literature review discussing mainly latest references (recent 5 years) is missing. Please share 5-10 existing works and summarize their performance and limitations.
AA: Thanks for this comment as it allows us to enlarge the general background on the topic. We have found several interesting recent contributions related to our new approach as the make usage of extensions of fuzzy sets applied to solve machine problems. Specifically, we have included the following new references • We have added a final sentence in the first paragraph of the introduction to point out the following recent applications of intuitionistic fuzzy sets

[15] T. Pencheva, M. Angelova, K. Atanassov, Genetic algorithms quality assessment implementing intuitionistic Fuzzy logic, Vol. 3-4, 2015 (2015).
[16] T. Pencheva, M. Angelova, Intuitionistic fuzzy logic implementation to assess purposeful model parameters genesis, Studies in Computational Intelligence 657 (2017) 179–203 (2017).
[17] T. Kim, E. Sotirova, A. Shannon, V. Atanassova, K. Atanassov, L.-C. Jang, Interval Valued Intuitionistic Fuzzy Evaluations for Analysis of a Student's Knowledge in University e-Learning Courses, International Journal of Fuzzy Logic and Intelligent Systems 18 (3) (2018) 190–195 (2018).
[18] C. Zhang, Classification rule mining algorithm combining intuitionistic fuzzy rough sets and genetic algorithm, International Journal of Fuzzy Systems 22 (5) (2020) 1694–1715 (2020).

- We have updated the list of references of applications of interval-valued or interval type 2 fuzzy sets in classification and control problems

[6] A. Tellez-Velazquez, H. Molina-Lozano, L. A. Villa-Vargas, R. Cruz-Barbosa, E. Lugo-Gonzalez, I. Z. Batyrshin, I. J. Rudas, A Feasible Genetic Optimization Strategy for Parametric Interval Type-2 Fuzzy Logic Systems, International Journal of Fuzzy Systems 20 (1) (2018) 318–338 (2018).
[8] T.-C. Lu, Genetic-algorithm-based type reduction algorithm for intervaltype-2 fuzzy logic controllers, Engineering Applications of Artificial Intelligence 42 (2015) 36–44 (2015).
[12] J. Carlos Guzman, I. Miramontes, P. Melin, G. Prado-Arechiga, Optimal Genetic Design of Type-1 and Interval Type-2 Fuzzy Systems for Blood Pressure Level Classification, Axioms 8 (1) (2019).

- We have included a new reference in the list of references of the usefulness of GAs to optimize this type of systems (first paragraph of Section 3 in page 8)

[46] T. Zhao, Y. Xiang, S. Dian, R. Guo, S. Li, Hierarchical interval type-2 fuzzy path planning based on genetic optimization, Journal of Intelligent and Fuzzy Systems 39 (1) (2020) 937–948 (2020).

- We have also included two related papers on developments on the basis of FARC-HD, as it is the base of IVFARC. The best configuration of the classifier in [25] is included in the experimental comparison to better support the quality of our new algorithm.

[25] G. Lucca, G. Dimuro, J. Fernandez, H. Bustince, B. Bedregal, J. Sanz, Improving the performance of fuzzy rule-based classification systems based on a non averaging generalization of CC-integrals named cF1F2-integrals, IEEE Transactions on Fuzzy Systems 27 (1) (2019) 124–134 (2019).
[67] G. Lucca, J. Antonio Sanz, G. Dimuro, B. Bedregal, H. Bustince, R. Mesiar, CF-integrals: A new family of pre-aggregation functions with application to fuzzy rule-based classification systems, Information Sciences 435 (2018)94–110 (2018).

RV:

4. Section 1, please summarize the contributions of the paper, preferably in point-form.

AA:

We have modified some sentences of the introduction section so that the contributions of the paper can be clearly observed. Specifically, we have rewritten both part of the fourth paragraph of the introduction (last paragraph of page 2) and the next one (the previous fifth paragraph, in the middle of the current page 3) so that the stages of the algorithm are stated in point-form. This way, the contributions of the paper are clearer.

RV:

5. Section 2.1, please elaborate on the reasons for the selection of [25] and [26].

AA:

In [25] (current reference 33) authors defined a method to generate linear orders based on aggregation functions. This way, one can construct easily different linear orders by using different aggregation functions. A particular case of this construction is the classical linear order defined by Xu and Yager in [26] (current reference 34). This latter order works fine in classification problems and in fact, we have submitted a paper to the IEEE TFS journal (it is accepted with minor changes) where we analyze the relationship among the operation used to perform de interval matching degree and the linear order. I attach below a table of this new paper where we present in rows different options to compute the interval matching degree and in columns we show three well known linear orders (lexicographical 1 and 2 as well as that of Xu and Yager, XY). The value in each cell is the averaged performance over 31 datasets. We are not going to describe all the matching degree options but the one corresponding to the product (the one we use in IVFARC) is the first one (REP-Prod), where we can see that the averaged performance of the linear order defined by Xu and Yager is better than that of the lexicographical ones. In fact, we can observe that the linear order defined by Xu and Yager provides the best performance for many options.

Method	$\leq_{Lex1}$	$\leq_{XY}$	$\leq_{Lex2}$
REP-Prod	78.96	79.67	79.17
REP-Min	78.92	79.52	79.51
REP-Hp	79.08	79.34	79.35
REP-Ob	79.00	79.83	79.33
REP-Gm	79.14	79.48	79.57
REP-Hm	79.19	79.39	79.41
ADM-Prod	78.88	79.14	79.11
ADM-Min	79.07	79.39	79.52
ADM-Hp	79.03	79.47	79.65
ADM-Ob	**79.32**	79.57	79.61
ADM-Gm	78.93	**79.93**	**79.75**
ADM-Hm	79.08	79.23	78.84

RV:

6. Section 2.1, authors mentioned "A typical operation among all types of FSs is the conjunction, which is usually modeled by t-norms [27] although new ways of aggregation interval information could be applied [28].". Please clarify what the approach in [28] is. Also, why [27] was chosen?

AA:

[27] (current reference 35) has been chosen as it is a book where aggregation functions are reviewed from the very beginning and you can find a good introduction to this theory. In [28] (current reference 36) authors presented a new way of aggregating interval information by means of the Choquet integral so that the relationship among the data to be aggregated can be taken into account. The usage of the Choquet integral implies the need of sorting the data, which are intervals in this case, so authors also tackle this problem. We have modified the paragraph so that the contribution in [28] (current reference 36) is briefly stated (third paragraph of page 5).

RV:

7. Section 3, paragraph 1, please specify the evolutionary computation as in "we have made usage of evolutionary computation,…".

AA:

Thanks for this comment as it allows us to clarify the details of our proposal. We have added a sentence at the end of the first paragraph of Section 3 (page 8) in order to specify that we use a GA to carry out the learning of the interval-valued fuzzy partitions and then, for each possible solution, we apply a differential evolution process to optimize the IVFRBCS learned when using each solution of the GA.

RV:

8. Section 3.1, paragraph 1, please address:

(i) Please explain the selection of triangular membership function.

(ii) Figure 2, please include x-label and y-label. Also, increase the font size of the variables.

AA:

We have selected a homogeneous partition of the input space using triangular shaped membership functions as it is recommended for promoting explainable artificial intelligence [1], which is focused on providing more explainable, interpretable, and transparent systems. Specifically, this process is desirable to increase the comprehensibility of the database [2].

[1] T. Miller, “Explanation in artificial intelligence: Insights from the social sciences,” Artificial Intelligence, vol. 267, pp. 1–38, 2019.

[2] J. Casillas, O. Cordón, F. H. Triguero, and L. Magdalena, Interpretability issues in fuzzy modeling. Springer, 2013, vol. 128.

For this reason, we have modified the sentence of the first paragraph of section 3.1 (page 8) to state these facts.

We have also included the x and y labels to figure 2.

RV:

9. Section 3.2, please address:

(i) Authors mentioned "The chromosome evaluation of our new method is slow and for this reason, we need to apply as much selective pressure as possible in order to be able to obtain good results in few generations. We have to stress that we use the global tuning of the semantics approach [38].", elaboration is needed because accuracy, divergence, convergence, time complexity could be considered and they may have been conflicting to each other.

(ii) What is/are the objective function(s)?

(iii) What is the fitness function?

AA:

We completely agree with the reviewer because accuracy and convergence are usually conflicting, since a fast convergence may imply reaching a local optimum and therefore, the obtained solution may not be as accurate as possible. However, in our case, the evaluation of a single chromosome requires a time demanding process because it is necessary to learn an IVFRBCS using the Apriori algorithm, which is slow when having many variables and classes, and optimizing it applying a differential evolution process. For this reason, we have sacrificed a possible increase of the accuracy of the system in order to obtain a solution in a reasonable period of time, which in turn it is also accurate. As pointed out, one of our future research lines would be to try to speed up the learning of the system for the sake of diminishing the selective pressure to avoid local optima and trying to obtain an increase on the accuracy of the system.

Regarding the objective /fitness function, it is the accuracy of the system, that is, the percentage of correctly classified examples. This way, we try to increase as much as possible the system's accuracy in a short period.

According to these comments, we have modified both the second paragraph of section 3.2 (page 10) to elaborate more on the accuracy/converge problem and the second item of this section (also in page 10) to clarify the fitness function used.

RV:

10. Table 1, do some of the datasets are imbalanced datasets?

AA:

This a really interesting comment because, as the reviewer imagines, some of the datasets have imbalanced distributions and this is one of our future research lines. In fact, we are currently working on this topic but using numerical fuzzy sets and binary imbalanced classification problems. We have submitted a new paper to the IEEE TFS journal and we are planning to expand the new methodology presented in that paper to work with IVFS so that we can study whether this type of fuzzy set is suitable for this framework or not. Furthermore, it would be interesting to study the effectiveness of these systems in imbalanced multi-class problems. According to these facts, we have pointed out this future research line in the last paragraph of Section 6 (page 28).

RV:

11. Section 4, paragraph 2, authors are suggested to cite the following articles as examples of 5-fold cross-validation.

(i) Predicting at-risk university students in a virtual learning environment via a machine learning algorithm

(ii) Reliable Accuracy Estimates from k-fold Cross Validation

AA:

Thanks for these references. We have added the second one as it is highly related to our research and it will be interesting for future readers as well.

RV:

12. It can be seen from Section 5.1 that authors have shared the analysis of IVFARC, the ideas of considerations are good. I have some comments:

(i) Divide into serval more subsections 5.1.1, 5.1.2, … for better organization.

(ii) Regarding the analysis of different settings, authors are suggested to show more scenarios (equal step size) for better illustration.

13. Section 5.1 and 5.2, authors are suggested to highlight clearly that what performance indicators are being selected to evaluate the performance of the methods. Also, revise the titles of tables to indicate what parameters are being considered.

14. Authors mentioned in early discussion that "The chromosome evaluation of our new method is slow…", section 5 should carry out analysis on this aspect.

15. Performance comparison between proposed work and latest existing works (recent 5 years) should be presented.

16. Authors have evaluated the performance of algorithms using many datasets, please elaborate in which conditions the proposed algorithms work better.

AA:

Thanks for these comments as they helped us to improve the analysis of the behavior of our new classifier. We have fused all of them in a single answer as all these comments are related to the experimental study. We have made the following modifications according to them:

- We have divided Section 1 in four sub-sections (5.1.1 to 5.1.4) to study four different scenarios as the reviewer has pointed out. This manner, the study is easier to understand.
- We have specified in the caption of Tables 3 and 6 that the performance is evaluated using the accuracy rate. In Table 9 we have specified that we show the average number of rules and antecedents (in parenthesis of the different classifiers) and in Table 10 (it is new in this version), we have specified that we measure the run-rime in seconds. This way, the performance metric used in each table is clear just looking at the caption.

The configuration of the different classifiers is specified in Section 4.1 (Table 2) and we have not added them into the caption of the tables as it may imply having too long captions.

- We have divided Section 5.2 in two sub-sections in order to also analyze the run-time of the algorithms as required by the reviewer (Section 5.2.2). This way, it can be observed (Table 11) that our new approach is slow when learning the system but after that it can be applied very fast as the classification of new examples is as fast as the remainder methods (milliseconds).
- We have added a new state-of-the-art fuzzy classifier that has been published recently and it is named C_{F1F2} in the comparison carried out in Section 5.1.4. This new method is closely related to our new classifier as it shares the learning process (using just numerical fuzzy sets) but it considers a new inference process where a generalization of the Choquet integral is used to fuse in a better manner the local information given by the fired rules (see the new reference [25] for details of the new considered classifier). Consequently, we have modified both Tables 6, 7 and 9 to include its results and the text of Section 5.1.4 and 5.2.1 to analyze them as well as Section 4.1 to introduce its set-up.
- The study of the conditions of the datasets where our new approach works better was made in Section 5.3. Anyway, in order to highlight the purpose of this section we have modified its title and we think it is clearer now. In this section, we can observe that IVFARC works better when the problems are more difficult according to different data complexity measures. This way, the hypothesis that IVFSs are needed when the definition of the membership functions is difficult is corroborated, since it would be the case in these problems as they have larger degrees of non-linearity.

RV:

17. Conclusion, authors have briefly presented the limitations and future work directions, please elaborate.

AA:

We have modified the last paragraph of the conclusions to add more future lines of research that can be applied to the current work so that the performance of IVFARC can be improved in terms of run-time, which is the main limitation of the new approach, and performance. Moreover, we have also pointed out another scenario where the new classifier can be applied like imbalance classification problems thanks to a previous reviewer's comment. Specifically, we plan to improve the systems run-time by trying to develop new methods to construct the initial IVFPs that do not need to use an evolutionary algorithm as well as introducing recent developments in fuzzy reasoning methods that make usage of the Choquet integral and its generalizations to fuse the information given by the fired rules.

REVIEWER 2

RV:

The paper is interesting and mathematically correct (I had not seen any serious mistakes). On the other hand, it is well-known that the intuitionistic fuzzy sets are equipollent with interval-valued fuzzy sets. Similar of the present research, but with apparatus of intuitionistic fuzzy sets are published by M. Angelova, K. Atanassov, O. Roeva, T. Pencheva, A. Shannon and other authors. So, the authors of the present paper must give some comparisons with already published research.

AA:

Thanks for this comment as it allows us to enlarge the general background on the topic. We have found several interesting recent contributions related to our new approach as the make usage of extensions of fuzzy sets applied to solve machine problems. Specifically, we have included the following new references

- We have added a final sentence in the first paragraph of the introduction to point out the following recent applications of intuitionistic fuzzy sets according to the reviewer's suggestion

  [15] T. Pencheva, M. Angelova, K. Atanassov, Genetic algorithms quality assessment implementing intuitionistic Fuzzy logic, Vol. 3-4, 2015 (2015).
  [16] T. Pencheva, M. Angelova, Intuitionistic fuzzy logic implementation to assess purposeful model parameters genesis, Studies in Computational Intelligence 657 (2017) 179–203 (2017).
  [17] T. Kim, E. Sotirova, A. Shannon, V. Atanassova, K. Atanassov, L.-C. Jang, Interval Valued Intuitionistic Fuzzy Evaluations for Analysis of a Student's Knowledge in University e-Learning Courses, International Journal of Fuzzy Logic and Intelligent Systems 18 (3) (2018) 190–195 (2018).
  [18] C. Zhang, Classification rule mining algorithm combining intuitionistic fuzzy rough sets and genetic algorithm, International Journal of Fuzzy Systems 22 (5) (2020) 1694–1715 (2020).

- We have updated the list of references of applications of interval-valued or interval type 2 fuzzy sets in classification and control problems

  [6] A. Tellez-Velazquez, H. Molina-Lozano, L. A. Villa-Vargas, R. Cruz-Barbosa, E. Lugo-Gonzalez, I. Z. Batyrshin, I. J. Rudas, A Feasible Genetic Optimization Strategy for Parametric Interval Type-2 Fuzzy Logic Systems, International Journal of Fuzzy Systems 20 (1) (2018) 318–338 (2018).
  [8] T.-C. Lu, Genetic-algorithm-based type reduction algorithm for intervaltype-2 fuzzy logic controllers, Engineering Applications of Artificial Intelligence 42 (2015) 36–44 (2015).
  [12] J. Carlos Guzman, I. Miramontes, P. Melin, G. Prado-Arechiga, Optimal Genetic Design of Type-1 and Interval Type-2 Fuzzy Systems for Blood Pressure Level Classification, Axioms 8 (1) (2019).

- We have included a new reference in the list of references of the usefulness of GAs to optimize this type of systems (first paragraph of Section 3 in page 8)

[46] T. Zhao, Y. Xiang, S. Dian, R. Guo, S. Li, Hierarchical interval type-2 fuzzy path planning based on genetic optimization, Journal of Intelligent and Fuzzy Systems 39 (1) (2020) 937–948 (2020).

- We have also included two related papers on developments on the basis of FARC-HD, as it is the base of IVFARC. The best configuration of the classifier in [25] is included in the experimental comparison to better support the quality of our new algorithm.

[25] G. Lucca, G. Dimuro, J. Fernandez, H. Bustince, B. Bedregal, J. Sanz, Improving the performance of fuzzy rule-based classification systems based on a non averaging generalization of CC-integrals named cF1F2-integrals, IEEE Transactions on Fuzzy Systems 27 (1) (2019) 124–134 (2019).
[67] G. Lucca, J. Antonio Sanz, G. Dimuro, B. Bedregal, H. Bustince, R. Mesiar, CF-integrals: A new family of pre-aggregation functions with application to fuzzy rule-based classification systems, Information Sciences 435 (2018)94–110 (2018).

REVIEWER 3

RV: The authors propose a learning approach for interval-valued fuzzy rule-based classification systems. It continues their past well-acknowledged papers in the field. The theory is correct and the experimental results are illustrative. The paper is definitely acceptable. I suggest its further improvement in terms of these comments that I hope it will improve its impact: 1. Please try, if possible, to replace "method" with "algorithm".
AA: Thanks for these comments. Regarding the first comment, we have replaced method with algorithm throughout the paper according to the reviewer's suggestion.

RV: 2. In this regard, please point out the steps of the algorithm.
AA: Thanks for this comment as it allows us to improve the description of our new approach. We have rewritten the previous fifth paragraph of the introduction, in the middle of the current page 3, so that the stages of the algorithm are stated in point-form. Furthermore, we have added some sentences at the end of the first paragraph of Section 3 (page 8) to describe the main steps of the proposed algorithm, which are graphically shown in Figure 1. This way, the steps of the algorithm are clearer in the current version.

RV: 3. Besides, authors are advised to cite a few papers to enlarge the general background and current status of fuzzy and nonlinear modeling. For example, the next few papers that produced important results in the field could be considered: Neural network based feature extraction for Assamese character and numeral recognition, International Journal of Artificial Intelligence, vol. 2, no. S9, pp. 37-56, 2009; New results in modelling derived from Bayesian filtering, Knowledge-Based Systems, vol. 23, no. 2, pp. 182-194, 2010; Medical image retrieval using vector quantization and fuzzy S-tree, Journal of Medical Systems, vol. 41, no. 18, pp. 1-16, 2017; ECG classification using three-level fusion of different feature descriptors, Expert Systems with Applications, vol. 114, pp. 54-64, 2018.
AA: Thanks for this comment as it allows us to enlarge the general background on extensions of fuzzy sets applied to solve nonlinear problems. Unfortunately, the

references provided by the reviewer do not use extensions of fuzzy sets so we have not included them in the manuscript. However, in our new search we have found very interesting related papers that have been included in the new version. Specifically

- We have added a final sentence in the first paragraph of the introduction to point out the following recent applications of intuitionistic fuzzy sets

  [15] T. Pencheva, M. Angelova, K. Atanassov, Genetic algorithms quality assessment implementing intuitionistic Fuzzy logic, Vol. 3-4, 2015 (2015).
  [16] T. Pencheva, M. Angelova, Intuitionistic fuzzy logic implementation to assess purposeful model parameters genesis, Studies in Computational Intelligence 657 (2017) 179–203 (2017).
  [17] T. Kim, E. Sotirova, A. Shannon, V. Atanassova, K. Atanassov, L.-C. Jang, Interval Valued Intuitionistic Fuzzy Evaluations for Analysis of a Student's Knowledge in University e-Learning Courses, International Journal of Fuzzy Logic and Intelligent Systems 18 (3) (2018) 190–195 (2018).
  [18] C. Zhang, Classification rule mining algorithm combining intuitionistic fuzzy rough sets and genetic algorithm, International Journal of Fuzzy Systems 22 (5) (2020) 1694–1715 (2020).

- We have updated the list of references of applications of interval-valued or interval type 2 fuzzy sets in classification and control problems

  [6] A. Tellez-Velazquez, H. Molina-Lozano, L. A. Villa-Vargas, R. Cruz-Barbosa, E. Lugo-Gonzalez, I. Z. Batyrshin, I. J. Rudas, A Feasible Genetic Optimization Strategy for Parametric Interval Type-2 Fuzzy Logic Systems, International Journal of Fuzzy Systems 20 (1) (2018) 318–338 (2018).
  [8] T.-C. Lu, Genetic-algorithm-based type reduction algorithm for intervaltype-2 fuzzy logic controllers, Engineering Applications of Artificial Intelligence 42 (2015) 36–44 (2015).
  [12] J. Carlos Guzman, I. Miramontes, P. Melin, G. Prado-Arechiga, Optimal Genetic Design of Type-1 and Interval Type-2 Fuzzy Systems for Blood Pressure Level Classification, Axioms 8 (1) (2019).

- We have included a new reference in the list of references of the usefulness of GAs to optimize this type of systems (first paragraph of Section 3 in page 8)

  [46] T. Zhao, Y. Xiang, S. Dian, R. Guo, S. Li, Hierarchical interval type-2 fuzzy path planning based on genetic optimization, Journal of Intelligent and Fuzzy Systems 39 (1) (2020) 937–948 (2020).

- We have also included two related papers on developments on the basis of FARC-HD, as it is the base of IVFARC. The best configuration of the classifier in [25] is included in the experimental comparison to better support the quality of our new algorithm.

  [25] G. Lucca, G. Dimuro, J. Fernandez, H. Bustince, B. Bedregal, J. Sanz, Improving the performance of fuzzy rule-based classification systems based on a non averaging generalization of CC-integrals named cF1F2-integrals, IEEE Transactions on Fuzzy Systems 27 (1) (2019) 124–134 (2019).
  [67] G. Lucca, J. Antonio Sanz, G. Dimuro, B. Bedregal, H. Bustince, R. Mesiar, CF-integrals: A new family of pre-aggregation functions with application to fuzzy rule-based classification systems, Information Sciences 435 (2018)94–110 (2018).

RV:

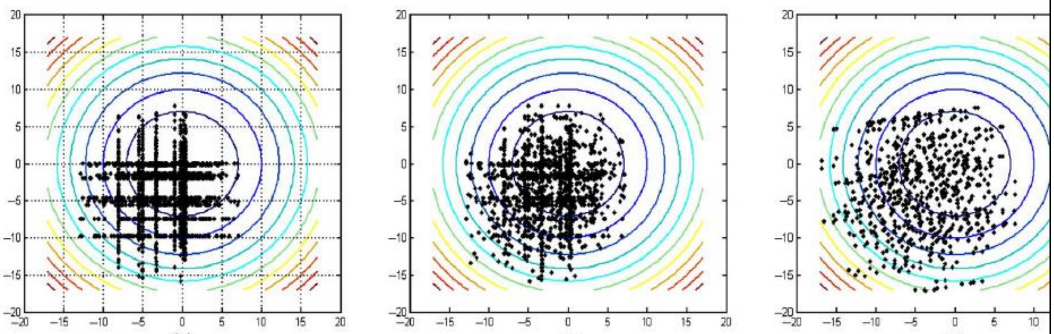
4. I did not find the optimization problem. Maybe it is not the case. But, if possible, its connection to the algorithm would be useful.

5. You could better highlight the effects of random parameters.

AA:

We have added a sentence in the first paragraph of section 3 (page 7) so that the optimization problem is clearly stated. Specifically, the learning of the IVFRBCS can be seen as an optimization problem where the best location and shape of each IVFS has to be found. In this manner, the objective of the optimization process is to provide the best location and shape of the IVFSs implying the learning of the most accurate classifier. This is the reason why the accuracy rate is used as the fitness function of both the GA and de DE processes. With the current modification we think that the purpose of the new algorithm is clearer than in the previous version of the manuscript.

Regarding the randomness of the parameters, this is the main reason why we have analyzed the effect of the probability of the application of the crossover operator, since the greater the probability the larger the randomness of the search is. This fact is depicted in the figure we attach below where the effect of the crossover probability in a two-dimensional optimization process according to the three sub-figures is shown: (a) the probability is 0 and the new solutions are created in a grid shape manner; (b) the probability is 0.5 and new solutions are more spread than in the previous scenario and (c) the probability is 1.0 and it can be observed that the new solutions are very spread. Therefore, it can be observed that the larger the probability is the more the randomness of new solution becomes. In the experimental study, we can observe that when using 0.9 as the crossover probability the results are not as good as those obtained when using 0.15 because the randomness degree is too much and the optimization process does get lost.



(a) (b) (c)

RV:
6. Giving a link to the programs and datasets in Section 5 will increase the area of readers.
AA: Thanks for this comment. We have added a link to the webpage where the datasets can be downloaded in a footnote in page 14 and we have pointed out in the footnote of page 17 that the code of IVFARC will be available in the web page where the results are located when the paper is published.

RV:
7. Please check the highlights to fulfill the length constraint.
AA: Thanks for the comment. We have modified the highlights so that they fulfil the journal’s requirements and we have highlighted the minimum percentage of improvement of our method according to the reviewer’s 1 comment.

*Highlights (for review)

- A new wrapper methodology to learn interval-valued fuzzy rule-based classifiers.
- Interval-valued fuzzy partitions are built such that the led to an accurate system.
- Antecedents of the rules use the most proper type of fuzzy set.
- State-of-the-art fuzzy classifiers are enhanced by at least 0.66% in average.
- IVFARC achieves good results when the problems are difficult.

A wrapper methodology to learn interval-valued fuzzy rule-based classification systems

Jos Antonio Sanz[a,*], Humberto Bustince[a]

[a]*Departamento of Estadstica, Informtica y Matemticas and Institute of Smart Cities, Universidad Publica de Navarra, Navarra, 31006 Spain*

Abstract

Learning an interval-valued fuzzy rule-based classification system is a challenge as its success directly depends on the interval-valued fuzzy partition used. In fact, the learning of an interval-valued fuzzy system usually starts by creating a partition composed of numerical fuzzy sets, which are used to build an initial fuzzy classifier. Then, it is augmented with interval-valued fuzzy sets whose shape is subsequently optimized to improve the system's performance. However, as in this methodology the fuzzy rules are learned using numerical fuzzy sets, the benefits of the interval-valued fuzzy sets may not be fully exploited. In this paper we define a new learning methodology that avoids building the initial fuzzy classifier but directly learns interval-valued fuzzy rules. To do so, we define a wrapper methodology to learn the interval-valued fuzzy partitions such that they lead to an interval-valued fuzzy rule-based classification system as accurate as possible. Moreover, our new method allows one to represent each membership function using the most proper type of fuzzy set for the sake of modeling the uncertainty in the best possible manner. Consequently, the antecedents of the rules can be formed of only numerical fuzzy sets, only interval-valued fuzzy sets or a mixture of both. The quality of the proposal is compared versus four state-of-the-art fuzzy classifiers like FARC-HD, IVTURS, FURIA and FARC-HD using an inference based on a generalization of the Choquet integral. We also compare our new approach besides its numerical fuzzy counterpart to clearly show the benefits of the usage of interval-valued fuzzy sets. Specifically, the average accuracy rate of our new method is 81.17%, which is at least 0.66% better than the remainder

*Corresponding author telephone number: +34 948169265
Email addresses: `joseantonio.sanz@unavarra.es` (Jos Antonio Sanz), `bustince@unavarra.es` (Humberto Bustince)

state-of-the-art fuzzy classifiers.

Keywords: Classification problems, Evolutionary fuzzy systems, Fuzzy rule-based classification systems, Interval-valued fuzzy sets, Interval type-2 fuzzy sets

1. Introduction

Interval Valued-Fuzzy Sets (IVFSs) [1] are an extension of numerical fuzzy sets (nFSs) [2]. Their usage can be beneficial when tackling difficult problems, that is, when the definition of the membership functions is complicated and IVFSs can offer an advantage when handling with the inherent uncertainty of this step [3]. In this sense, the length of the interval can represent the lack of knowledge of the expert when defining the nFSs [4]. Due to their advantages they have been widely used to tackle machine learning problems like regression [5, 6, 7, 8, 9], clustering [10, 11] and classification [12, 13, 14]. Atanassov's intuitionistic fuzzy sets are closely related to IVFSs and they have been successfully applied to tackle real-world problems [15, 16, 17, 18].

Among the previously mentioned applications, addressing classification problems using IVFSs has received less attention from the research community. There are two major approaches to build an interval-valued fuzzy system [19]:

- To generate an initial fuzzy system, which uses nFSs, and augmenting it with IVFSs afterwards.
- To develop a method that directly learns an interval-valued fuzzy system.

Most of the approaches to learn Interval-Valued Fuzzy Rule-based Classification Systems (IVFRBCSs) [20] are based on the former option whereas few of them follow the latter and they are mostly based on clustering [21].

In order to directly learn an IVFRBCS a key question is how to define the Interval-Valued Fuzzy Partition (IVFP) for each variable of the problem. Moreover, one can wonder which is the most proper type of FS, nFSs or IVFSs, used to model each linguistic label composing those partitions. Consequently, we can think of the linguistic labels modeling as a learning problem, where the best membership functions need to be constructed.

The main contribution of this paper is the definition of a new wrapper-based methodology whose main aim is to learn directly an accurate IVFRBCS, that is, without using an initial FRBCS and augmenting it subsequently. To achieve it, we have the following objectives:

- To define an evolutionary process to learn the IVFPs, where a chromosome represents the join of all the IVFPs of the problem. In the evaluation, the IVFPs represented by a chromosome are used to construct an initial IVFRBCS, which is also optimized, and its training performance is used as the fitness value. In this manner, we learn the IVFPs leading to an accurate IVFRBCS.

- To propose a learning approach that builds directly interval-valued fuzzy rules, which is necessary to apply our new methodology.

- To produce interpretable fuzzy rules whose antecedents can combine nFSs and IVFSs. That is, they use IVFSs when they are really needed and nFSs otherwise.

Specifically, since the new methodology is a wrapper, the stages of the algorithm are:

1. We first use a Genetic Algorithm (GA) [22] to learn the shape of each IVFS, which includes the lateral position, the amplitude of the support of both the lower and the upper bounds as well as the amplitude of the core of the upper bound. The key component of the GA is the evaluation, where we learn an IVFRBCS whose training accuracy rate is used as fitness value. In this paper, we extend on IVFSs the two first stages of the FARC-HD classifier [23], since it is an state-of-the-art fuzzy classifier, but it could be another learning approach that directly builds interval-valued fuzzy rules.
2. Next, the learned IVFRBCS is optimized by a differential evolutionary algorithm (DE) [24], whish is applied to tune the shape of the upper bound of the IVFSs and to perform a rule selection process.
3. At the end of the process, each fuzzy rule of the learned IVFRBCS can be composed of only nFSs, only IVFSs or a mixture of both. That is, IVFSs are used when they are appropriate and their shape is adjusted to the specific classification problem, which leads to a highly accurate fuzzy classifier.

To determine the quality of the new algorithm, the experimental study is first aimed at studying the best configuration of the wrapper method to learn IVFRBCSs. Moreover, we also analyze whether the usage of IVFSs is beneficial or not by comparing our new algorithm versus its numerical fuzzy counterpart. Finally, we also compare its behavior, in terms of its performance and rule base size, versus four state-of-the-art fuzzy classifiers like FARC-HD [23], FARC-HD using a generalization of the Choquet integral in the inference [25] and FURIA [26] as representatives of FRBCSs and IVTURS [27] as representative of IVFRBCSs. Moreover, we try to figure out when our new algorithm works better than the methods used in the comparison. The study is conducted using 29 numerical datasets selected from the KEEL dataset repository [28] and the results are supported by a proper statistical study [29, 30, 31]. In order to highlight the novelty of the approach, we show an example of an interval-valued fuzzy rule base (see Table 10 and Figure 3) learned by our new algorithm to show that the antecedents of the fuzzy rules are composed of different types of FSs (just nFS, just IVFSs or a mixture of both).

This paper is organized as follows. In Section 2, we present some related preliminary concepts that are necessary to understand the paper. In Section 3 we describe in detail our new wrapper-based methodology to learn IVFRBCSs. The experimental framework and the analysis of the obtained results are reported in Sections 4 and 5, respectively. In Section 6 we draw the main conclusions.

2. Preliminaries

In this section, we first review several preliminary concepts in both IVFSs (Section 2.1) and IVFRBCSs (Section 2.2).

2.1. Interval-Valued Fuzzy Sets

This section is aimed at introducing the theoretical concepts related to IVFSs necessary to understand the new method. Therefore, we start recalling the definition of IVFSs, whose history and relationship with other type of FSs as interval type-2 FSs can be found in [32].

Let $L([0,1])$ be the set of all closed subintervals in $[0,1]$:

$$L([0,1]) = \{\mathbf{x} = [\underline{x}, \overline{x}] | (\underline{x}, \overline{x}) \in [0,1]^2 \text{ and } \underline{x} \leq \overline{x}\}.$$

Definition 1. *[1] An interval-valued fuzzy set A on the universe $U \neq \emptyset$ is a mapping $A_{IV} : U \to L([0,1])$, so that*

$$A_{IV}(u_i) = [\underline{A}(u_i), \overline{A}(u_i)] \in L([0,1]), \textit{ for all } u_i \in U.$$

Obviously, $[\underline{A}(u_i), \overline{A}(u_i)]$ is the interval membership degree of the element u_i to the IVFS A.

In our new method we need to compare interval membership degrees and consequently, we need a total order relationship for intervals using admissible orders [33]. Specifically, we have selected the usage of that defined by Xu and Yager[1] in [34], which is based on the score and accuracy degrees as shown in (1).

$$\begin{aligned}&[\underline{x},\overline{x}] \leq [\underline{y},\overline{y}] \text{ if and only if}\\ &\underline{x}+\overline{x} < \underline{y}+\overline{y} \text{ or } \underline{x}+\overline{x} = \underline{y}+\overline{y} \text{ and } \overline{x}-\underline{x} \geq \overline{y}-\underline{y}\end{aligned} \tag{1}$$

Using (1) it is easy to observe that $0_L = [0,0]$ is the smallest element in $L([0,1])$. We have to stress that when we need to compare an interval and a constant value k, we create the corresponding interval for the value, that is, $[k,k]$.

The antecedents of the fuzzy rules are modeled by IVFSs or nFSs. A typical operation among all types of FSs is the conjunction, which is usually modeled by t-norms [35] although new ways of aggregation interval information could be applied, like the usage of the Choquet integral to take into account the relationship among the interval data to be aggregated [36]. In this paper, we model the intersection by means of t-representable interval-valued t-norms[2] [37] without zero divisors, that is, they verify that $\mathbf{T}(\mathbf{x},\mathbf{y}) = 0_L$ if and only if $\mathbf{x} = 0_L$ or $\mathbf{y} = 0_L$. We denote them $\mathbf{T}_{T_a,T_b}$, since they are represented by T_a and T_b, which are the t-norms applied over the lower and the upper bounds, respectively. That is, $\mathbf{T}_{T_a,T_b}(\mathbf{x},\mathbf{y}) = [\mathbf{T_a}(\underline{\mathbf{x}},\underline{\mathbf{y}}),\mathbf{T_b}(\overline{\mathbf{x}},\overline{\mathbf{y}})]$.

Finally, we present the interval arithmetical operations [38] we apply. Let $[\underline{x},\overline{x}]$, $[\underline{y},\overline{y}]$ be two intervals in $\mathbb{R}^+$ so that $\mathbf{x} \leq_L \mathbf{y}$, the rules of interval arithmetic we use are:

- Addition: $[\underline{x},\overline{x}] + [\underline{y},\overline{y}] = [\underline{x}+\underline{y},\overline{x}+\overline{y}]$.
- Subtraction: $[\underline{x},\overline{x}] - [\underline{y},\overline{y}] = [|\underline{y}-\overline{x}|,\overline{y}-\underline{x}]$.
- Multiplication: $[\underline{x},\overline{x}] * [\underline{y},\overline{y}] = [\underline{x}*\underline{y},\overline{x}*\overline{y}]$.
- Division: $\frac{[\underline{x},\overline{x}]}{[\underline{y},\overline{y}]} = [\min(\min(\frac{\underline{x}}{\underline{y}},\frac{\overline{x}}{\overline{y}}),1),\min(\max(\frac{\underline{x}}{\underline{y}},\frac{\overline{x}}{\overline{y}}),1)]$ with $\underline{y} \neq 0$.
- Division by a constant k, with $k > 1$: $[\frac{\underline{x}}{k},\frac{\overline{x}}{k}]$.

[1]This order is a specific case of those presented in [33] when using $\alpha = 0.5$.

[2]A nFS is an IVFS whose lower and upper membership functions are equal.

2.2. Interval-Valued Fuzzy Rule-Based Classification Systems

Solving a classification problem consists in learning a mapping function called *classifier* from a set of training examples, named *training set*, that allows new examples to be classified. The training set is composed of P examples, $x_p = (x_{p1}, \ldots, x_{pn}, y_p)$, where x_{pi} is the value of the i-th attribute $(i = 1, 2, \ldots, n)$ of the p-th training example. Each example belongs to a class $y_p \in \mathbb{C} = \{C_1, C_2, ..., C_m\}$, where m is the number of classes of the problem.

FRBCSs are a widely used technique to deal with classification problems [39]. When using this type of system, each of the n attributes is described by a set of linguistic terms modeled by their corresponding membership functions. Consequently, they provide an interpretable model as the antecedent part of the fuzzy rules is composed of a subset of these linguistic terms as shown in (2).

$$\text{Rule } R_j : \text{ If } x_1 \text{ is } A_{j1} \text{ and } \ldots \text{ and } x_n \text{ is } A_{jn} \text{ then Class} = C_j \text{ with } RW_j \tag{2}$$

where R_j is the label of the jth rule, $x = (x_1, \ldots, x_n)$ is an n-dimensional pattern vector, A_{ji} is an antecedent fuzzy set representing a linguistic term, C_j is the class label, and RW_j is the rule weight [40].

An IVFRBCS [20] is an extension of a FRBCS where the linguistic terms, A_{ji}, of the fuzzy rules are modeled using IVFSs instead of nFSs. A widely applied methodology to create them is to execute a fuzzy rule learning algorithm and, once the fuzzy rules have been learned, the nFSs are extended into IVFSs, whose shapes are subsequently optimized for the sake of improving the performance of the system [14].

One of the key components of an IVFRBCS is the interval inference mechanism, which uses the L interval-valued fuzzy rules composing the model to determine the class of previously unknown examples. Specifically, the steps of the Interval-Valued Fuzzy Reasoning Method (IV-FRM) [27] when classifying an example $x_p = (x_{p1}, \ldots, x_{pn})$ are the following ones:

1. *Interval matching degree*: It quantifies the strength of activation of the if-part for all rules (L) in the system with the example x_p:

$$\begin{aligned} &[\underline{\mu_{A_j}}(x_p), \overline{\mu_{A_j}}(x_p)] = \mathbf{T}_{T_a,T_b}([\underline{\mu_{A_{j1}}}(x_{p1}), \overline{\mu_{A_{j1}}}(x_{p1})], \\ &\ldots, [\underline{\mu_{A_{jn}}}(x_{pn}), \overline{\mu_{A_{jn}}}(x_{pn})]), \qquad j = 1, \ldots, L. \end{aligned} \tag{3}$$

2. *Interval association degree*: for each rule, the interval matching degree is weighted by its rule weight:

$$[\underline{b_j}(x_p), \overline{b_j}(x_p)] = [\underline{\mu_{A_j}}(x_p), \overline{\mu_{A_j}}(x_p)] * [\underline{RW_j}, \overline{RW_j}] \quad j = 1, \dots, L. \tag{4}$$

3. *Interval pattern classification soundness degree for all classes*. The positive interval association degrees are aggregated by class applying an aggregation function *f*.

$$\begin{aligned} [\underline{Y_k}, \overline{Y_k}] &= f_{R_j \in RB;\, C_j = k}([\underline{b_j}(x_p), \overline{b_j}(x_p)] | \\ &[\underline{b_j}(x_p), \overline{b_j}(x_p)] > 0_L), \qquad k = 1, \dots, m. \end{aligned} \tag{5}$$

4. *Classification*. A decision function F is applied over the interval soundness degrees:

$$F([\underline{Y_1}, \overline{Y_1}], ..., [\underline{Y_m}, \overline{Y_m}]) = \text{arg} \max_{k=1,...,m}([\underline{Y_k}, \overline{Y_k}]) \tag{6}$$

The above IV-FRM is the one used in IVTURS [27] but without applying IV-REFs [41] to speed up the inference. We have to point out that when every linguistic term is modeled by nFSs (the lower and the upper membership degrees are the same) the IV-FRM is the same as the classical numerical FRM [42].

3. A Wrapper Methodology to Learn Interval-Valued Fuzzy Rule-based Classification Systems

This section is aimed at describing in detail our new wrapper-based methodology for generating an IVFRBCS, whose flowchart es shown in Figure 1. This new algorithm learns directly interval-valued fuzzy rules, that is, it does not apply a fuzzy rule learning algorithm and once the fuzzy rules are learned they are extended into IVFSs. When learning an IVFRBCS one of the main problems is the definition of the IVFSs because if they are not properly defined (shape and position) the quality of the final IVFRBCS will not be good. Therefore, this problem can be seen as an optimization problem, where it is necessary to learn the best IVFSs for the tackled problem. That is, to optimize (learn) the position and shape of each IVFS so that they lead to an accurate IVFRBCS. To cope with it, we propose a wrapper methodology to be able to achieve two goals at the same time: 1) to build good IVFSs and 2) to produce an accurate IVFRBCS. To do it, we have made usage of evolutionary computation, since it is a recommended tool to optimize type-2 systems [43] and it has also proved to work properly for this task [44, 45, 46]. Specifically, we can observe in Figure 1 that we use a GA to learn the IVFPs and then, for each chromosome, we apply a DE to optimize the learned IVFRBCS using the IVFP represented by that solution. Finally, the best solution (IVFRBCS) will be used to classify new testing examples.

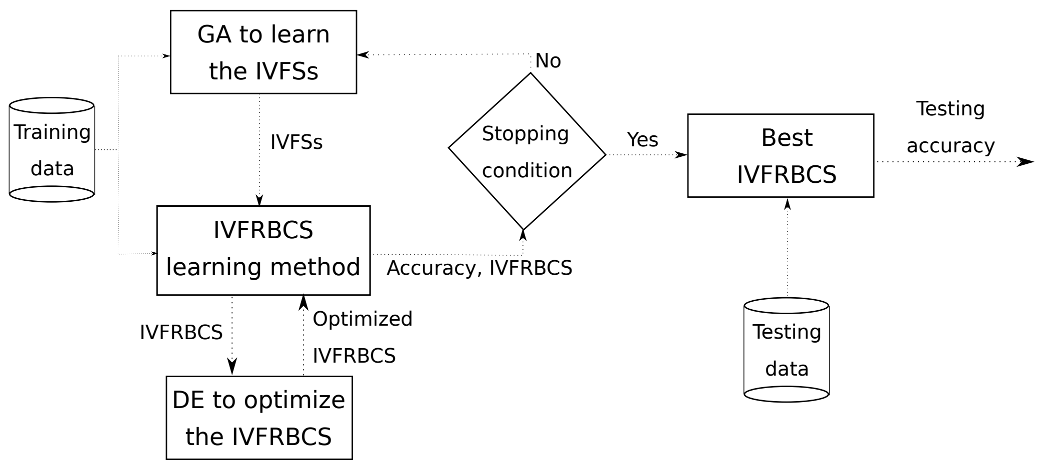


Figure 1: Flowchart of the new learning algorithm of IVFRBCSs.

In the remainder of this section we describe in detail all the sub-components of the system starting from the modeling and construction of the IVFSs (Section 3.1). Then, we introduce the GA (Section 3.2) that wraps the learning algorithm of IVFRBCSs (Section 3.3).

3.1. Interval-Valued Fuzzy Sets Generation and Parametrization

IVFSs are modeled by the lower and the upper membership functions. In order to construct them, we start creating the lower bound by performing an homogeneous partition of the input space using triangular shaped membership functions as shown in dotted lines in Figure 2, which is desirable to increase the comprehensibility of the database [47] to promote the explainable artificial intelligence [48]. As a result, the support of the lower bounds (S_L) is determined[3].

Next, for each IVFS, we add the upper bound, which is modeled by a trapezoidal shaped membership function that is symmetrically centered around the

[3] A different shape could be used by defining the proper parameters and changing accordingly the search space in the optimization stages.

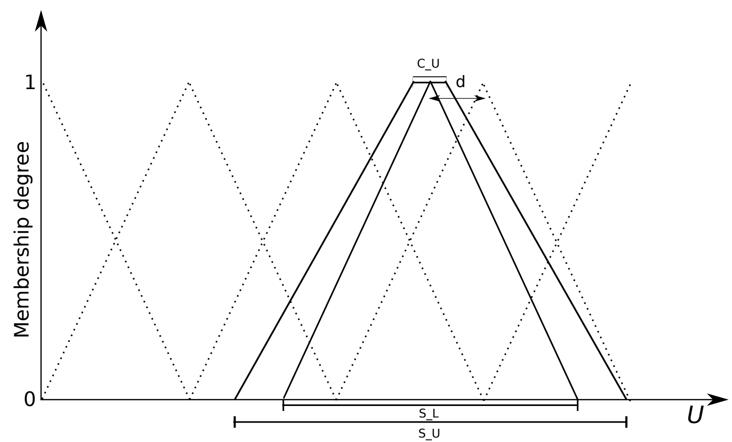


Figure 2: Example of the shape and representation of an IVFS.

lower membership function. Consequently, its shape is modeled by two parameters as can be observed in the membership function shown with a solid line in Fig. 2:

- The amplitude of the support, which is represented by the parameter S_U. This amplitude can vary from being equal to that of the lower bound (less support) to being twice than that of the lower bound (largest support).
- The amplitude of the core, which is represented by the parameter C_U. The largest core equals the amplitude of support of the lower membership function and the less one is when it collapses in the center of the lower bound.

Consequently, when using the less amplitude of both the support and the core of the upper bound we obtain a nFS and in the remainder situations an IVFS is generated.

Finally, for each IVFS, we obtain its final position by displacing it to the left or to the right according to the 2-tuples model [49]. This displacement is represented by the parameter d in Fig. 2 (parameter α in [49]), where the original IVFS would be located around the membership function in the right of the solid IVFS. As in any proposal for displacing linguistic labels the displacement is bound at half the support of the lower bound of the neighboring labels.

3.2. Genetic Algorithm to Learn the IVFSs leading to an Accurate IVFRBCS

In this section we present the evolutionary process used to wrap the learning of the IVFSs in such a way that they lead to obtaining an accurate IVFRBCS. As we have presented in the previous section, IVFSs are modelled by means of three parameters: d, S_U and C_U (the value for the parameter S_L is automatically obtained when performing the homogeneous partition of the input space). Consequently, the aim of the evolutionary process is to learn the most proper values of these parameters for each IVFS.

Among all the evolutionary techniques we have considered GAs [22] because they allow one to control the selective pressure. The chromosome evaluation of our new algorithm is slow and for this reason, we need to apply as much selective pressure as possible in order to be able to obtain good results in few generations. That is, we look for a fast convergence by sacrificing a possible increase of the system's accuracy in order to obtain a solution in a reasonable period of time. We have to stress that we use the global tuning of the semantics approach [49]. That is, a linguistic label is modeled by the same IVFS in all the fuzzy rules using it. Consequently, the interpretability of the system is maintained to a large degree as we will show in Section 5.2.

The specific components of the GA used to learn the shape of the IVFSs and the subsequent IVFRBCS are the following ones:

- **Coding scheme**. Each chromosome is composed of as many genes as the number of linguistic labels times three, since we are learning the parameters d, S_U and C_U. We use a real coding for all the genes, where their ranges and meanings are:
  - Parameter d: according to the 2-tuples model [49] the range is $[-0.5, 0.5]$, where -0.5 and 0.5 correspond to the left and right most displacements ($\pm\ \frac{S\_L}{2}$). The value 0.0 represents the initial position.
  - Parameters S_U and C_U: we use a gene for each one (g_{S_U} and g_{C_U}, respectively) encoded in the range $[0.0, 1.0]$, where 0.0 represents the less amplitude of the parameters (producing a nFS) and 1.0 represents the largest IVFS ($S\_U = 2 \cdot S\_L$ and $C\_U = S\_L$, respectively).

- **Chromosome Evaluation**. We evaluate the quality of each chromosome using the most common metric for classification, i.e. the accuracy rate, since it represents an IVFP that will be used to learn an IVFRBCS. To do it, we first generate the IVFSs represented by the chromosome, which involves placing them in the position determined by d and translating g_{S_U} and g_{C_U} into the real ranges of S_U and C_U, respectively. Then, using these IVFSs we learn an IVFRBCS, which can be also optimized, using the approach presented in Section 3.3. To alleviate the computational burden, we only optimize these IVFRBCSs candidates to improve the best performance. If the IVFRBCS is optimized the values of the genes are updated accordingly.

- **Population size**. We consider 10 individuals because the chromosome evaluation is time demanding.

- **Initial Gene Pool**. In the initial population we include three chromosomes representing these three situations: 1) IVFSs in their initial position, having a triangular shaped upper bound whose amplitude of the support is 50% larger than that of the lower bound (every gene for d and C_U is set to 0.0 and those for S_U are set to 0.5); 2) nFSs centered in the initial position (all the genes are set to 0.0) and 3) the largest IVFSs centered in the initial position (every gene for d is set to 0.0 and those for S_U and C_U are set to 1.0). The remainder chromosomes are randomly generated in the proper ranges.

- **Selection**. We need to apply as much selective pressure as possible because of the time demanding evaluation we apply. Therefore, we opt for a tournament with replacement ($k = 3$) so that it is likely to happen that a good individual wins the tournament.

- **Crossover Operator**. We cross two individuals, who generate two offspring, when the crossover probability is larger than a threshold (otherwise the parents become the offspring). Specifically, we apply the Parent Centrix BLX operator [50] as it provides a good trade-off between exploration and exploitation.

- **Mutation**. We apply a Gaussian mutation (mean equal to the gene value and standard deviation equal to 0.1) when the mutation probability is larger than a threshold. This probability is applied to know both the individuals and the genes to be mutated.

- **Replacement**. In order to increase the selective pressure we select the best individuals (as many as individuals in the population) between parents and offspring.

3.3. Interval-Valued Fuzzy Rule Learning Algorithm and Optimization

In this section we first present the approach to learn an IVFRBCS (Section 3.3.1) and then we also introduce an optimization process to refine it in order to improve the system's performance as much as possible (Section 3.3.2).

3.3.1. Interval-valued Fuzzy Rule Learning Algorithm

To apply our new wrapper-based methodology it is necessary to have a learning algorithm able to provide directly interval-valued fuzzy rules. For this reason, we modify the two first stages of the FARC-HD classifier[4] [23], since it is an state-of-the-art fuzzy classifier, in such a way that they work with IVFSs instead of with nFSs and as a result, they provide interval-valued fuzzy rules directly. Consequently, we name our new algorithm IVFARC because this learning algorithm will provide Interval-Valued Fuzzy Association Rules for Classification.

An association rule represents a dependency among items in a database using expression like $A \to B$, where A and B are sets of items and $A \cap B \neq \emptyset$ [51]. Association rules can be used to cope with classification problems [52], since their antecedent part can be composed of fuzzy terms and the consequent part can be formed of the predicted class label and the rule weight as shown in (2). When there is at least an IVFS in the antecedent of a fuzzy rule we have an interval-valued fuzzy rule.

In order to generate the rules, we apply the Apriori algorithm [53] where each item is a linguistic label (IVFSs) and the support and confidence degrees are computed using the following equations, which are adapted to work with IVFSs:

$$Supp(A_j) = \frac{\sum_{p=1}^{P} [\underline{\mu_{A_j}}(x_p), \overline{\mu_{A_j}}(x_p)]}{P} \tag{7}$$

$$Conf(A_j \to C_j) = \frac{\sum_{x_p \in Class C_j} [\underline{\mu_{A_j}}(x_p), \overline{\mu_{A_j}}(x_p)]}{\sum_{p=1}^{P} [\underline{\mu_{A_j}}(x_p), \overline{\mu_{A_j}}(x_p)]} \tag{8}$$

When the system is composed of only nFSs (their lower and upper bound are the same), (7) and (8) are the fuzzy support and confidence [54], respectively.

After applying the Apriori algorithm, a set of interval-valued fuzzy rules is obtained, which can be composed of a large number of rules. For this reason, as FARC-HD does, we apply a second stage to select the most promising rules among

[4] A different learning algorithm could be applied but it needs to provide an IVFRBCS directly.

this set applying the fuzzy pattern weighting scheme[5] defined in [23], which assigns a weight (w_p) to each example. Obviously, we also need to modify this method to work with intervals instead of numbers. Specifically, we only need to modify the equation used to measure the quality of each rule, $R_j : A_j \rightarrow C_j$, as follows:

$$wWRAcc''(R_j) = \frac{n''(A_j \cdot C_j)}{n'(C_j)} \cdot \left(\frac{n''(A_j \cdot C_j)}{n''(A_j)} - \frac{n(C_j)}{P} \right) \tag{9}$$

where

$$n''(A_j) = \sum_{p=1}^{P} [w_p, w_p] \cdot [\underline{\mu_{A_j}}(x_p), \overline{\mu_{A_j}}(x_p)]$$

$$n''(A_j \cdot C_j) = \sum_{x_p \in ClassC_j} [w_p, w_p] \cdot [\underline{\mu_{A_j}}(x_p), \overline{\mu_{A_j}}(x_p)]$$

$$n'(C_j) = \sum_{x_p \in ClassC_j} [w_p, w_p]$$

3.3.2. Differential Evolutionary Process to Optimize the IVFRBCS

According to our previous papers in the topic [20, 4, 56, 27], optimizing the shape of the upper bound of the IVFSs can lead to an improvement of the system's performance. Furthermore, in a fuzzy rule base we can find irrelevant, redundant, erroneous or conflicting rules, which may perturb the performance of the system [57]. Consequently, a rule reduction process is often applied with a double aim: 1) to improve the accuracy of the classifier by removing useless rules and 2) to ease the readability of the system.

Therefore, we propose to apply an optimization process in order to obtain both the most suitable shape for the upper bound of the IVFSs and a reduced set of rules cooperating properly among themselves. To do so, we consider the usage of the classical DE method [58] as it usually offers a fast convergence. In this method, for each target vector (individual) a test vector is obtained by applying differential mutation among three randomly selected vectors (with repetition) and the uniform crossover. Then, if the test vector is better than the target vector, the latter is replaced by the former. Consequently, the selective pressure is high as the search is guided by the best performing individuals implying a fast convergence

[5]The original pattern weighting scheme was introduced in [55] to evaluate the quality of rules in the APRIORI-SD algorithm.

of the algorithm, which is the reason why we have applied it. The specific features of the DE process for our optimization problem are:

- **Coding scheme**. The chromosome is composed of two parts, one for the parameters related to the shape of the upper membership function (S_U and C_U) and the other to perform the rule selection process. The first part uses a real coding, in the range $[0, 1]$ as explained in the GA, and it is composed of as many genes as twice the total number of linguistic labels. The second part considers a binary coding and it is composed of as many genes as the number of rules composing the model, where the values 1 and 0 implies the usage of the corresponding rule in the IV-FRM or not.

- **Initial pool**. We initialize four individuals, which use all the rules (those genes are set to 1), to represent the following situations: 1) the initial IVFRBCS (the values of the chromosome of the GA representing the IVFSs are copied); 2) the usage of nFSs (all the genes used to optimize S_U and C_U are set to 0.0; 3) the largest IVFSs (genes for S_U and C_U set to 1.0) and 4) a triangular shaped upper bound whose amplitude of the support is 50% larger than that of the lower bound (genes for S_U and C_U set to 0.5 and 0.0, respectively). The remainder individuals are generated at random in the ranges of the genes.

- **Fitness**. In order to measure the quality of the individuals we adapt the shape of the IVFSs, we recompute the rule weights and we measure the quality of the tuned IVFRBCS using the accuracy rate.

- **Stopping Criteria**. The search is stopped when either the maximum number of generations is reached or when 100% accuracy is obtained.

4. Experimental framework

We use 29 real world data-sets selected from the KEEL dataset repository [59][6]. All the datasets used in the study are numerical ones taking into account that in case of integer coding their number of different values is larger than the number of linguistic labels used per variable. Therefore, the use of fuzzy logic make sense in this scenario. We have to point out that the 19 first datasets shown in Table 1

[6] `http://www.keel.es/`

were used in [27], where IVTURS was defined. Table 1 summarizes the properties of these datasets, showing, for each dataset, the identifier (Id.) as well as the name (Dataset), the number of instances ($\#Inst$), the number of attributes ($\#Att$) and the number of classes ($\#Class$). Examples with missing values have been removed, e.g., in the *bands, cleveland* and *wisconsin* datasets.

Table 1: Datasets used in this study

Id.	Dataset	#Inst	#Att	#Class
Bal	Balance	625	4	3
Cle	Cleveland	297	13	5
Eco	Ecoli	336	7	8
Hab	Haberman	306	3	2
Ion	Ion	351	33	2
Iri	Iris	150	4	3
Mag	Magic	1,902	10	2
New	Newthyroid	215	5	3
Pag	Page-blocks	547	10	5
Pen	Penbased	1,099	16	10
Pim	Pima	768	8	2
Spe	Spectfheart	267	44	2
Tae	Tae	151	5	3
Tit	Titanic	2,201	3	2
Two	Twonorm	740	20	2
Veh	Vehicle	846	18	4
Win	Wine	178	13	3
WiR	Winequality-Red	1,599	11	11
Wis	Wisconsin	683	11	2
Ban	Banana	5,300	2	2
Bnd	Bands	365	19	2
Bup	Bupa	345	6	2
Gla	Glass	214	9	6
Pho	Phoneme	5,404	5	2
Rin	Ring	740	20	2
Sat	Satimage	643	36	7
Shu	Shuttle	5,800	9	7
Wdb	Wdbc	569	30	2
Yea	Yeast	1,484	8	10

To determine the performance of the classifiers, we have applied a 5-fold cross validation model [60]. The final results shown in this study are the average of the accuracy rate, which is the percentage of correctly classified examples, among the five testing folders.

In order to give statistical support to the analysis of the results we consider the usage of hypothesis validation techniques [61]. We use non-parametric tests as suggested in [29, 30, 31], where it is shown that their use in the field of machine learning is highly recommended. Specifically, we use the aligned Friedman test [62] besides the post-hoc Holm's test [63] for group comparisons and the Wilcoxon test [64] to perform pair-wise comparisons.

4.1. Methods Set-up

This section is aimed at introducing the configurations that have been considered for the different classifiers used along the experimental study, namely, the different versions of our new IVFRBCS (named IVFARC), the FARC-HD classifier [23], the IVTURS classifier [27], the FURIA algorithm [26] and FARC-HD using the generalization of the Choquet integral defined in [25] (named C_{F1F2}).

The values of the parameters of the classifiers are shown in Table 2. As there are approaches sharing the same stage, in the first column the algorithm(s) sharing the component pointed out in the second column are introduced, where the values considered for the parameters of the corresponding stage are shown in the last column.

Table 2: Set-up of the parameters of the classifiers.

Algorithm(s)	*Stage*	*Parameters*
FARC-HD IVTURS IVFARC C_{F1F2}	Rule learning	Linguistic labels per variable: 5 Rule weight: certainty factor Minimum Support: 0.05 Minimum Confidence: 0.8 Maximum depth: 3 Parameter k_t: 2
FARC-HD IVTURS C_{F1F2}	CHC	Evaluations: 20000 Individuals: 50 α parameter: 0.02 Bits per gen: 30
IVFARC	GA	Generations: 10 Individuals: 10 Tournament size in GA: 3 Crossover probability GA: 0.9 Mutation probability GA: 0.05
IVFARC	DE	Vectors: 3 Generations: $\{500, 1000\}$ Individuals: 10 Crossover probability: $\{0.15, 0.9\}$ Parameter F: 1.0
FURIA	Rule learning	Number of optimizations: 2 Linguistic labels per variable: 3

The parameters of the GA used in our new methodology are widely used in the literature except the low number both of generations and individuals, which will be analyzed in Section 5.1. Regarding the DE process, the number of individuals is the same than that of the GA and we want to assess the performance when varying both the number of generations and the crossover probability.

Finally, all the classifiers use the additive combination as the inference process except C_{F1F2} that uses a generalization of the Choquet integral. Specifically, we have considered the best pair of functions of the C_{F1F2} classifier according to [25]. The remainder of the set-up of the classifiers used in the comparison is that recommended by their authors.

5. Experimental Results

In this section we analyze the behavior of our new proposal, which is named IVFARC, in terms of performance, run-time and rule base size (we also show an example of a rule base). Specifically, the aims of the experimental study are:

- To determine the best configuration of the DE process carried out in IVFARC. Our base configuration is the one introduced in Table 2 considering 500 generations and a crossover probability of 0.15 ($IVFARC^{0.15}_{500}$). We want to analyze: 1) whether increasing the number of generations to 1,000 ($IVFARC^{0.15}_{1000}$) is beneficial or not; 2) if broaden the application of the crossover operator by setting the crossover probability to 0.9 ($IVFARC^{0.9}_{500}$) allows one to enhance the results.
- To check whether $IVFARC^{0.15}_{500}$ is able to converge or not.
- To analyze the effect of the usage of IVFSs by comparing our new proposal versus its numerical fuzzy counterpart ($IVFARC_{nFS}$).
- To compare IVFARC versus four state-of-the-art fuzzy classifiers, namely, FARC-HD [23], C_{F1F2} and IVTURS [27], since they share the same fuzzy rule learning algorithm, and FURIA [26].
- To study the features of the datasets were $IVFARC^{0.15}_{500}$ behaves well in order to try to understand why (or when) the algorithm is working fine.

5.1. Analyzing the Performance of IVFARC

The results provided in testing and the standard deviation ($\pm$) by the different settings of the new classifier are presented in Table 3[7]. In this table, classifiers are

[7]The detailed results both in training and testing as well as the rule base sizes, run-times and the data complexity measures can be queried at `https://github.com/JoseanSanz/IVFARC`. The code of the algorithm will be available in this site when the paper is published.

shown by columns and the datasets by rows, where the best result is highlighted in **bold-face**. We also show the averaged mean among all the datasets obtained in training by all the classifiers, Mean$_{Tr}$. Furthermore, we show a summary of the testing results by adding two rows where we compute the number of wins (Wins) and defeats (Loses) of the method in the column versus $IVFARC^{0.15}_{500}$. Finally, in the last column (AvgGen) we show the averaged generation of the GA where $IVFARC^{0.9}_{500}$ gets the best result.

Table 3: Accuracy rate in testing (standard deviation) by the different settings of the new approach. The best result is highlighted in **bold-face**. The last column shows the averaged generation (AvgGen) of the GA where $IVFARC^{0.15}_{500}$ gets the maximum.

Dataset	$IVFARC^{0.15}_{500}$	$IVFARC^{0.15}_{1000}$	$IVFARC^{0.9}_{500}$	$IVFARC_{nFS}$	AvgGen
Bal	**88.64±1.64**	87.20±2.12	85.28±1.21	84.80±2.33	1.00±0.00
Cle	**59.26±3.80**	57.58±2.16	58.25±3.56	57.24±2.26	1.00±0.00
Eco	**83.35±4.07**	80.66±3.09	82.44±2.21	81.26±4.70	4.80±4.26
Hab	71.55±6.28	67.97±3.76	70.91±6.80	**73.18±3.73**	3.80±3.49
Ion	89.18±4.80	89.75±4.86	89.75±3.94	**91.19±4.89**	1.00±0.00
Iri	95.33±3.80	94.67±3.80	**96.67±3.33**	95.33±3.80	1.60±1.20
Mag	79.91±2.30	**80.02±2.11**	78.23±2.54	79.70±3.18	3.20±2.99
New	96.74±2.65	**97.21±1.95**	95.81±2.55	95.35±3.68	2.60±1.50
Pag	**94.52±2.14**	93.60±2.53	94.34±1.22	**94.52±1.74**	2.00±2.00
Pen	92.36±1.74	**93.27±1.91**	91.55±2.37	91.27±1.42	1.00±0.00
Pim	**76.69±1.45**	76.43±3.25	75.39±2.51	75.91±2.78	4.60±4.03
Spe	79.01±5.29	**80.52±1.67**	76.80±4.71	79.76±3.74	7.60±1.96
Tae	**59.59±2.92**	58.97±7.05	56.97±4.49	57.63±5.24	6.00±3.10
Tit	**78.87±1.48**	**78.87±1.48**	**78.87±1.48**	**78.87±1.48**	1.00±0.00
Two	**92.84±2.37**	92.16±1.32	**92.84±1.95**	90.68±2.31	1.00±0.00
Veh	67.50±3.11	**68.21±2.79**	63.48±2.50	67.38±3.03	1.00±0.00
Win	96.05±4.77	94.94±3.06	95.49±1.58	**97.75±3.09**	1.00±0.00
WiR	59.54±2.40	59.04±3.11	**60.79±2.59**	58.66±2.53	5.60±3.20
Wis	96.34±1.38	96.78±0.83	**96.93±1.31**	95.76±1.57	2.60±2.24
Ban	88.30±1.60	**88.87±0.76**	84.77±1.77	84.45±0.83	5.80±2.71
Bnd	68.00±4.10	**70.78±7.13**	68.29±5.72	66.37±6.47	1.00±0.00
Bup	68.12±7.60	**71.30±5.16**	68.70±3.64	68.41±4.51	2.40±1.20
Gla	66.83±8.23	**68.70±2.96**	65.91±4.62	66.84±4.29	3.60±2.15
Pho	81.38±0.68	79.91±2.30	80.33±1.05	**82.22±1.34**	7.00±2.61
Rin	**92.16±2.51**	91.62±2.17	89.86±2.74	90.00±2.68	1.00±0.00
Sat	78.23±3.26	79.01±1.80	76.36±1.90	**79.94±2.49**	3.60±1.85
Shu	98.80±0.60	**99.40±0.35**	97.52±2.31	97.75±2.19	5.20±3.60
Wdb	**96.84±0.49**	96.31±2.00	**96.84±1.00**	95.78±1.14	3.00±2.28
Yea	**57.89±4.34**	56.13±4.62	56.87±3.52	57.48±3.05	6.00±1.90
Mean	**81.17±3.17**	81.03±2.83	80.22±2.80	80.53±2.98	
Mean$_{Tr}$	87.73±0.78	**88.22±0.77**	86.19±0.86	87.59±0.60	
Wins		13	6	8	
Loses		15	20	18	

In the remainder of this section we analyze the behaviour of IVFARC starting from the configuration of the DE process (Section 5.1.1). Then, we study the convergence of IVFARC (Section 5.1.2) and the suitability of the usage of IVFSs (Section 5.1.3). Finally, we compare our new algorithm versus four state-of-the-art fuzzy classifiers to support its quality (Section 5.1.4).

5.1.1. Determining the best set-up of the differential evolutionary process of IVFARC

In this section, we analyze the behaviour of the different settings of the DE process carried out in our new algorithm. Analyzing the results of the three configurations of the DE process we can observe that the best trade-off between the training and testing results is obtained by our base configuration ($IVFARC_{500}^{0.15}$). The approach using the largest number of evaluations, $IVFARC_{1000}^{0.15}$, enhances our base configuration in 13 cases whereas it provides a worse result in 15. However, $IVFARC_{1000}^{0.15}$ over-fits more than the base configuration. On the other hand, when increasing the crossover probability, it seems that the algorithm gets lost as the average training and testing results are worse than those of the base configuration, which is reflected in the number of loses versus $IVFARC_{500}^{0.15}$.

In order to provide an statistical support to the previous comments, we have carried out an Aligned Friedman and Holm tests to compare these three approaches. The p-value provided by the Aligned Friedman test is 1.50E-5, which implies the existence of statistical differences, and the ranks of the algorithms are shown in the second column of Table 4, where the method related to the less one is the best one. Consequently, our base configuration, $IVFARC_{500}^{0.15}$, is the best ranking one and consequently it is used as control method for the Holm's test, where the obtained APVs are shown in the last column of this table. From these results, we can observe that $IVFARC_{500}^{0.9}$ is outperformed whereas there are not statistical differences versus $IVFARC_{1000}^{0.15}$. All in all, we select our base configuration as the best algorithm because it is the best ranking configuration and it is the one obtaining the best generalization capability.

Table 4: Aligned Friedman and Holm tests to compare the different settings of IVFARC.

Method	Rank	APV
$IVFARC_{500}^{0.15}$	34.66	
$IVFARC_{1000}^{0.15}$	39.40	0.47
$IVFARC_{500}^{0.9}$	57.95	8.91E-4

5.1.2. Studying the convergence of IVFARC

We also want to figure out whether the best setting, $IVFARC^{0.15}_{500}$, is able to converge or not. This study is due to the fact that in this algorithm both the number of generations and the number of individuals of the GA is low (10), since the evaluation of a chromosome is time demanding. Therefore, using this setting the number of IVFRBCSs generated is limited to 100. Consequently, it is interesting to know if this setting is enough for its convergence.

The last column of Table 3 shows for each dataset the average generation among the 5 folders (besides the standard deviation, $\pm$) where the GA gets the maximum. We can observe that in almost all the cases the new algorithm arrives to its maximum performance before half of the generations. Therefore, we can conclude that thanks to the high selective pressure applied in the GA, $IVFARC^{0.15}_{500}$ is able to converge.

5.1.3. Analyzing the suitability of the usage of IVFSs

Once we have determined the best configuration of the DE process, we want to study the usefulness of the usage of IVFSs. To do so, we compare $IVFARC^{0.15}_{500}$ versus its numerical fuzzy counterpart ($IVFARC_{nFS}$). That is, the same algorithm but working only with nFSs, which implies changing the optimization process of both the GA and DE, as these sets are modeled by a less number of parameters. Specifically, in the GA we only optimize the parameter d (as S_U and C_U do not form part of nFSs) and in the DE we only accomplish the rule selection process. The set-up of the algorithm is the same than that of $IVFARC^{0.15}_{500}$ to make a fair comparison. Looking at the results shown in Table 3, we can observe that the usage of IVFSs implies a better generalization of the system and it allows one to enhance the testing results in 18 out of the 29 datasets of the study. In order to strengthen this conclusion, we have applied the Wilcoxon's statistical test to compare both approaches, whose results are shown in Table 5. In light of these results we can conclude that the usage of IVFSs implies an improvement of the system's performance with a high level of confidence.

Table 5: Wilcoxon Test to compare $IVFARC^{0.15}_{500}$ (R^{+}) versus its numerical fuzzy counterpart (R^{-}).

Comparison	R^+	R^-	p-value
$IVFARC^{0.15}_{500}$ vs. $IVFARC_{nFS}$	317.5	117.5	0.0325

5.1.4. Comparison against state-of-the-art fuzzy classifiers

Finally, we want to study whether the performance of our new algorithm is competitive versus four state-of-the-art fuzzy classifiers like FARC-HD, IVTURS, FURIA and FARC-HD using the novel FRM where a generalization of the Choquet integral is applied (C_{F1F2}). The results of these five classifiers are introduced in Table 6, which has the same structure as Table 3. In this case, we also show the average mean both in training and testing among all the datasets (All) as well as using only the 19 datasets common with the paper where IVTURS [27] was presented (Com.). We can observe that the best averaged result among these five classifiers in the two scenarios is the one provided by $IVFARC^{0.15}_{500}$. Finally, looking at the two scenarios, we can observe that IVTURS provides a very competitive performance using the common datasets (and in a lesser degree C_{F1F2}) but their behaviour is not that good using all the datasets whereas FURIA presents the opposed behaviour.

On the one hand, analyzing the classifiers based on the same fuzzy rule learning approach (FARC-HD, IVTURS and C_{F1F2}) our new proposal is obtaining the best result on two thirds of the datasets considered in the study. In order to support its superiority, we have applied the Aligned Friedman's test in the two scenarios showing the results in Table 7, where we can see that $IVFARC^{0.15}_{500}$ is the best ranking method in the two cases. The p-values are 2.58E-5 and 0.0014 using all the datasets and the common ones, respectively. Consequently, we have carried out the post-hoc Holm's test to compare these classifiers. The obtained results are introduced in Table 7, where we can observe statistical differences in favor to our new algorithm using all the datasets. However, as we have observed in Table 6, when using the common datasets IVTURS and C_{F1F2} are very competitive and there are not statistical differences.

On the other hand, when comparing $IVFARC^{0.15}_{500}$ with FURIA, we can observe that although our new algorithm obtains a better averaged mean, the number of datasets where each classifier provides the best result is similar. This fact is confirmed by the Wilcoxon test, whose results are shown in Table 8, as there are not statistical differences between them. However, when using only the common datasets, we can observe that $IVFARC^{0.15}_{500}$ outperforms FURIA as IVTURS did in [27].

5.2. Studying the Rule Base and Run-time

This section is aimed at analyzing the size of the rule base generated by the different algorithms and we also show an example of a rule base generated by $IVFARC^{0.9}_{500}$ (Section 5.2.1). Furthermore, we study the run-time of the classi-

Table 6: Accuracy rate in testing by the different state-of-the-art classifiers. The best result is highlighted in **bold-face**.

	Dataset	$IVFARC_{500}^{0.15}$	FARC-HD	IVTURS	FURIA	C_{F1F2}
	Bal	**88.64±1.64**	85.76±1.04	85.76±1.04	83.68±2.81	86.24±2.43
	Cle	59.26±3.80	55.88±2.80	**59.60±6.18**	56.57±3.35	56.23±3.49
	Eco	**83.35±4.07**	80.07±2.64	78.58±3.65	80.06±3.25	81.26±3.78
	Hab	71.55±6.28	71.22±5.63	**72.85±8.64**	72.55±5.36	72.53±5.46
	Ion	89.18±4.80	88.89±6.18	**92.89±3.27**	89.75±2.26	89.75±3.91
	Iri	95.33±3.80	94.00±5.48	**96.00±2.79**	94.00±5.96	**96.00±4.35**
	Mag	79.91±2.30	**81.12±2.53**	79.76±3.37	80.65±0.96	79.18±2.97
	New	**96.74±2.65**	96.28±3.89	95.35±2.33	94.88±3.03	96.28±3.53
	Pag	94.52±2.14	94.34±1.52	95.07±1.55	95.25±1.20	**95.43±2.16**
	Pen	92.36±1.74	**93.18±1.82**	92.18±2.83	92.45±2.33	92.00±2.89
	Pim	**76.69±1.45**	74.87±2.50	75.90±2.65	76.17±2.70	75.00±2.89
	Spe	79.01±5.29	77.53±3.96	**80.52±3.15**	77.88±8.12	78.65±5.74
	Tae	**59.59±2.92**	56.28±6.48	57.68±9.96	45.61±9.16	56.32±8.37
	Tit	**78.87±1.48**	**78.87±1.48**	**78.87±1.48**	78.51±1.13	**78.87± 1.48**
	Two	**92.84±2.37**	89.19±2.98	92.30±0.60	88.11±2.73	92.16±3.08
	Veh	67.50±3.11	68.32±3.16	67.38±2.37	**70.21±2.58**	67.97±0.98
	Win	96.05±4.77	96.60±3.10	**97.19±3.96**	93.78±3.77	96.62±2.36
	WiR	59.54±2.40	**60.91±1.97**	58.28±1.81	59.72±2.78	59.35±2.96
	Wis	96.34±1.38	96.34±0.53	96.49±1.09	**96.63±0.83**	**96.63±0.40**
	Ban	88.30±1.60	85.79±0.87	81.70±1.09	**88.57±0.90**	84.19±1.68
	Bnd	68.00±4.10	69.39±4.74	67.70±2.08	69.40±5.34	**69.68±6.69**
	Bup	68.12±7.60	67.25±2.43	67.54±5.29	**70.14±6.03**	63.77±4.35
	Gla	66.83±8.23	64.04±6.35	67.31±4.18	**72.91±7.02**	68.72±6.98
	Pho	81.38±0.68	80.92±1.35	80.00±1.35	**85.90±0.69**	80.83±1.22
	Rin	**92.16±2.51**	91.08±2.50	87.57±3.30	85.54±4.24	90.41±3.22
	Sat	78.23±3.26	80.71±1.86	75.90±1.03	**82.27±1.74**	78.38±3.22
	Shu	98.80±0.60	95.36±2.41	91.82±2.34	**99.68±0.13**	95.86±0.40
	Wdb	**96.84±0.49**	95.96±1.00	95.61±1.75	95.78±2.19	94.91±2.43
	Yea	57.89±4.34	**58.56±0.98**	55.86±2.41	58.22±2.36	57.75±2.64
Com.	Mean	**81.96±3.07**	81.03±3.14	81.72±3.30	80.34±3.39	81.39±3.34
Com.	Mean_{Tr}	88.99±0.72	**89.77±0.67**	87.80±0.74	86.35±1.58	89.44±0.79
All	Mean	**81.17±3.17**	80.30±2.90	80.13±3.02	80.51±3.27	80.38±3.31
All	Mean_{Tr}	87.73±0.78	**88.44±0.83**	85.72±0.81	87.20±1.64	88.05±0.77
All	Wins		8	8	16	10
All	Loses		19	19	13	18

Table 7: Aligned Friedman and Holm tests to compare $IVFARC_{500}^{0.15}$, FARC-HD, IVTURS and C_{F1F2}.

Method	All		Com.	
	Rank	APV	Rank	APV
$IVFARC_{500}^{0.15}$	44.21		31.42	
C_{F1F2}	61.62	0.07	40.16	0.45
FARC-HD	63.07	0.07	49.21	0.04
IVTURS	65.10	0.05	33.21	0.80

Table 8: Wilcoxon Test to compare $IVFARC^{0.15}_{500}$ (R^+) versus FURIA (R^-).

Datasets	Comparison	R^+	R^-	p-value
All	$IVFARC^{0.15}_{500}$ vs. FURIA	241	194	0.6114
Com.	$IVFARC^{0.15}_{500}$ vs. FURIA[8]	139	51	0.0766

fiers in Section 5.2.2.

5.2.1. Analyzing the Rule Base Size and Showing an Example

In Table 9 we present the rule base size of the different classifiers. Specifically, we show the averaged number of rules and antecedents per rule (number in brackets) in the 29 datasets composing the study. As can be observed, all the algorithms sharing the same fuzzy rule learning method (all the methods except FURIA) provide rule bases with similar sizes. Among them, the largest rule base is provided by $IVFARC^{0.9}_{500}$ possibly due to the fact that it gets lost in the DE search process and the smallest one is provided by FARC-HD, which is understandable as its fitness function considers the minimization of the number of rules. On the other hand, FURIA obtains a less number of fuzzy rules, which is possible mainly due to two facts: 1) the generated fuzzy rules do not need to cover all the input space because it uses an rule stretching mechanism to classify uncovered examples and 2) the same linguistic term is not modeled by the same membership function in all the fuzzy rules, which implies a larger degree of flexibility in the definition of the cell representing each rule at the cost of obtaining a system with a less global interpretability.

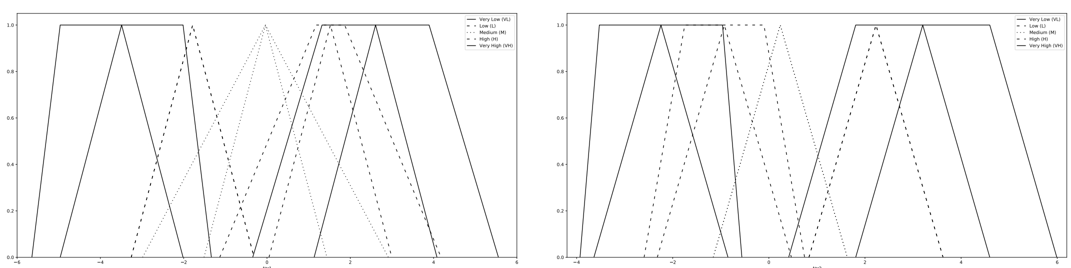
Table 9: Averaged number of rules (and antecedents) among all the datasets.

Method	RB size
$IVFARC^{0.15}_{500}$	26.52 (2.26)
$IVFARC^{0.15}_{1000}$	26.65 (2.27)
$IVFARC^{0.9}_{500}$	29.95 (2.29)
$IVFARC_{nFS}$	25.55 (2.16)
FARC-HD	23.86 (2.11)
IVTURS	25.61 (2.09)
FURIA	14.59 (2.79)
C_{F1F2}	28.29 (2.16)

In order to show an example of the rule base learned by IVFARC we use the first fold of the banana dataset. Table 10 shows the fuzzy rules leaned for this problem, whose membership functions are depicted in Figure 3. It can be seen that the rules can be composed of: 1) only nFSs (rules 3 and 4); only IVFSs (rules 1, 2, 5, 7, 8 and 12) and 3) a mixture of nFSs and IVFSs (rules 6, 9, 10 and 11).

Table 10: Rule base generated in the first fold of the banana dataset.

Rule ID	Rule
1	Att2 IS VH: -1.0 CF: [0.672, 0.732]
2	Att1 IS M: -1.0 CF: [0.571, 0.598]
3	Att1 IS L AND Att2 IS H: -1.0 CF: [0.984, 0.984]
4	Att1 IS L AND Att2 IS M: -1.0 CF: [0.415, 0.415]
5	Att1 IS M AND Att2 IS VL: -1.0 CF: [0.754, 0.931]
6	Att1 IS H AND Att2 IS M: -1.0 CF: [0.574, 0.711]
7	Att1 IS VH: 1.0 CF: [0.445, 0.572]
8	Att1 IS M: 1.0 CF: [0.402, 0.429]
9	Att1 IS L AND Att2 IS L: 1.0 CF: [0.736, 0.744]
10	Att1 IS M AND Att2 IS M: 1.0 CF: [0.570, 0.656]
11	Att1 IS H AND Att2 IS H: 1.0 CF: [0.464, 0.596]
12	Att1 IS H AND Att2 IS VL: 1.0 CF: [0.377, 0.799]



(a) Attribute 1 (Att 1) (b) Attribute 2 (Att 2)

Figure 3: IVFSs used in the rule base shown in Table 10.

We must point out that the interpretability of the system is maintained as we apply a global tuning of semantics approach in the evolutionary process. That is, the membership function used to model a linguistic term is the same in all the rules using it. The novelty of the algorithm is that each label is modelled using the most proper type of fuzzy set in order to improve the handing of the uncertainty inherent to their definition, which implies a leap on the system's performance.

5.2.2. *Study of the Run-time*

In order to study the run-time of the algorithms, we have measured (in seconds) the time required by all the approaches to learn the fuzzy rules and to classify all the testing examples. The averaged times among all the datasets of the study are shown in the second and third columns of Table 11, where we can observe the following facts: 1) all the methods require a few milliseconds to classify all the testing examples; 2) the learning run-time of the approaches derived from our new methodology, as expected, are the largest ones. Obviously, the largest is the one provided by the approach using the DE process applying a larger number of generations and the fastest approach is the one using nFSs as the search space is smaller; 3) the learning run-time of the methods using a single evolutionary process, specially FARC-HD, is much less than those derived from our new method and 4) the fastest method to build the system is FURIA as it does not include an evolutionary method to perform the optimization task. In light of these findings, we can conclude that all the approaches considered in this study are able to work in real-time environments as long as the learning stage can be performed off-line and the testing phase is really fast.

Table 11: Averaged run-time (in seconds) among all the datasets.

Method	Run-time	
	Learning	Testing
$IVFARC_{500}^{0.15}$	3363.03	0.0089
$IVFARC_{1000}^{0.15}$	4815.72	0.0091
$IVFARC_{500}^{0.9}$	3405.88	0.0088
$IVFARC_{nFS}$	498.08	0.0036
FARC-HD	16.94	0.0044
IVTURS	1614.40	0.0480
FURIA	0.97	0.0001
C_{F1F2}	68.33	0.0124

5.3. *Under What Conditions Does IVFARC Work Well?*

In this section we try to understand when the usage of our new approach provides good results. To do it we apply data complexity measures [65, 66] in order to study if IVFARC works better than the methods used in this paper ($IVFARC_{nFS}$, FARC-HD, IVTURS, FURIA and C_{F1F2}) according to certain features of the selected datasets. Specifically, we have computed the data complexity measures that are valid for multi-class classification problems, namely, F3, N1, N2, N3, N4 and T1[9].

Table 12 graphically shows the comparison of IVFARC versus the methods sharing the learning process, namely, FARC-HD, IVTURS, $IVFARC_{nFS}$ and

[9] We have discarded T2 as it does not differentiate by classes and its relevance is less obvious [65].

C_{F1F2} according to pairs of data complexity measures whereas in Figure 4 we show the comparison versus FURIA. In both figures we show using a diamond and a dot those datasets where IVFARC is better and worse than the other method, respectively. We can observe that IVFARC seems to works better than the four methods sharing the learning approach (Table 12) when F3 is low and N4 is high. We recall that F3 and N4 represent the maximum individual feature efficiency and the non-linearity of the nearest neighbor classifier (1NN). That is, the less the value of F3 (and the larger that of N4) the more difficult the problem is. Consequently, IVFARC is working better than these four methods when the problem is difficult and therefore, the definition of the membership functions plays a key role and the usage of the most proper type of fuzzy set for each label can imply an increase on the system's performance. On the other hand, the comparison between IVFARC and FURIA (Figure 4) does not provide a clear answer but it seems to be a trend where IVFARC works fine when F3 is low and N2 (ratio of average intra/interclass nearest neighbor distance) is high, that is, there are also difficult problems.

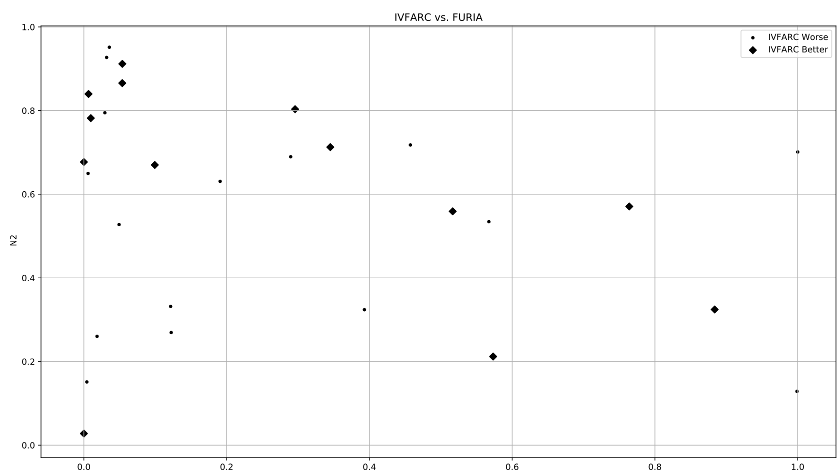
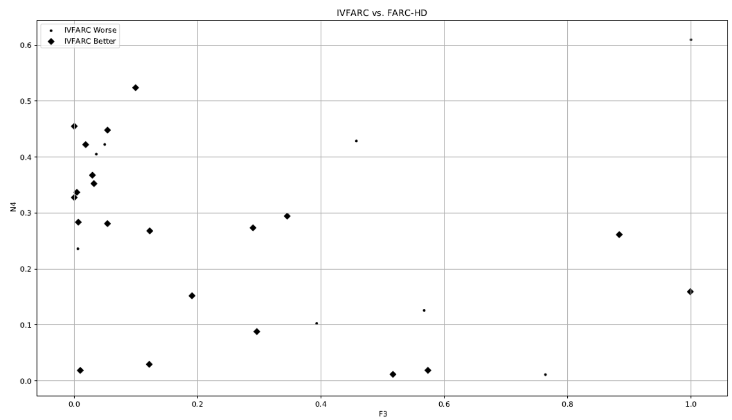


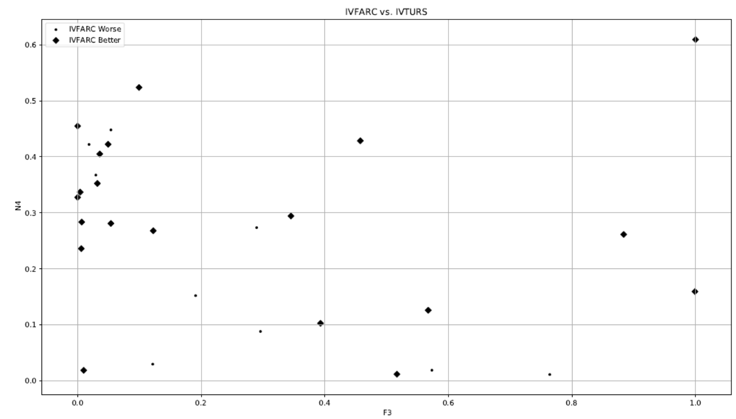
Figure 4: Comparison of IVFARC versus FURIA according to F3 and N2.

6. Conclusion

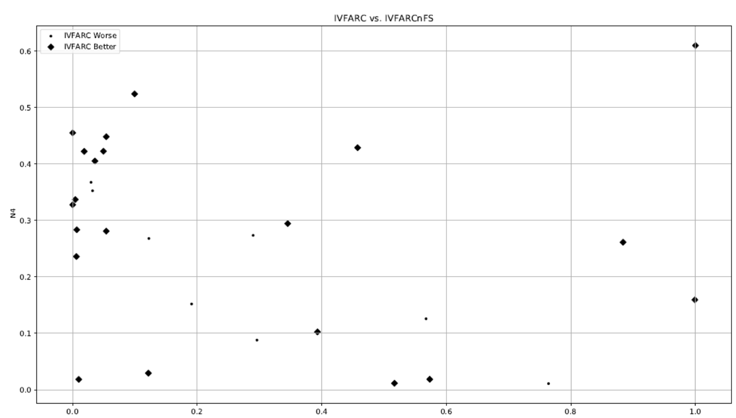
In this paper we have proposed a new learning algorithm of IVFRBCSs. Specifically, our proposal does not use an initial fuzzy system that is subsequently aug-



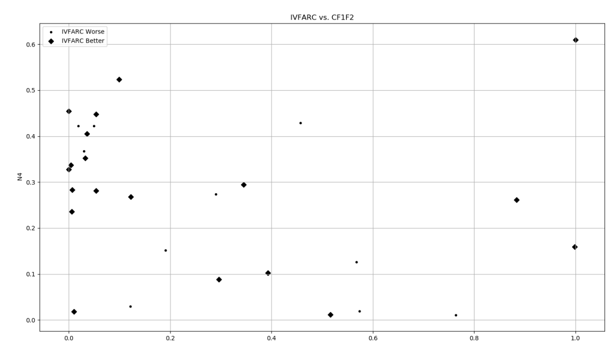
(a) IVFARC vs. FARC-HD



(b) IVFARC vs. IVTURS



(c) IVFARC vs. $IVFARC_{nFS}$



(d) IVFARC vs. C_{F1F2}

Table 12: Comparison of IVFARC versus fuzzy classifiers sharing the learning process according to F3 and N4.

mented with IVFSs but it directly learns interval-valued fuzzy rules. Moreover, the antecedent of the rules can be composed of nFSs, IVFSs or a mixture of both. To do so, we have wrapped the learning of the IVFPs in a GA, where to evaluate them an IVFRBCS is built and optimized applying a DE process. Consequently, the new algorithm not only learns the best IVFPs but also the best IVFRBCS. Our new approach is named IVFARC as the IVFRBCS learning algorithm is based on fuzzy association rules for classification.

The quality of IVFARC has been empirically tested in an study carried out over 29 datasets. From the experimental study we can highlight the following facts: 1) the usage of IVFSs is beneficial as our method outperforms its numerical fuzzy counterpart; 2) IVFARC allows one to enhance the result provided by two related state-of-the-art fuzzy classifiers like FARC-HD and IVTURS whereas its performance is comparable to that of FURIA; 3) the rule base size is similar to those of the related fuzzy classifiers but the antecedents of the fuzzy rules use the most proper kind of fuzzy set, which leads to an improvement of the performance; 4) the interpretability of the system is maintained and 5) the new algorithm works better than the other classifiers used in the study when the problems are difficult according to data complexity measures, which can be due to the method developed to learn the IVFPs.

The main limitation of the algorithm is its run-time as the new classifier is composed of two evolutionary processes and consequently, the speed up of the algorithm will be tackled in the near future in order to make the method applicable in a wider set of problems. Specifically, the learning of the IVFPs without the usage of evolutionary computation would help to achieve it. Moreover, we also want to study whether the generalizations of the Choquet integral developed in [25, 67] can be extended to work with IVFSs and consequently, they could imply a leap on the performance of IVFARC. Finally, it would be interesting to tackle imbalanced classification problems using IVFRBCSs to check if they can improve the results of the current fuzzy classifiers.

Acknowledgments

This work was supported in part by the Spanish Ministry of Science and Technology (project PID2019-108392GB-I00 (AEI/UE, FEDER))

References

[1] R. Sambuc, Function ϕ-flous, application a l'aide au diagnostic en pathologie thyroidienne, Ph.D. thesis, University of Marseille (1975).

[2] L. A. Zadeh, Fuzzy sets, Information and control 8 (3) (1965) 338–353.

[3] J. M. Mendel, Advances in type-2 fuzzy sets and systems, Information Sciences 177 (1) (2007) 84–110.

[4] J. Sanz, A. Fernández, 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, International Journal of Approximate Reasoning 52 (6) (2011) 751–766.

[5] K. Naik, C. Gupta, E. Fernandez, Design and implementation of interval type-2 fuzzy logic-pi based adaptive controller for dfig based wind energy system, International Journal of Electrical Power and Energy Systems 115. doi:10.1016/j.ijepes.2019.105468.

[6] A. Tellez-Velazquez, H. Molina-Lozano, L. A. Villa-Vargas, R. Cruz-Barbosa, E. Lugo-Gonzalez, I. Z. Batyrshin, I. J. Rudas, A Feasible Genetic Optimization Strategy for Parametric Interval Type-2 Fuzzy Logic Systems, International Journal of Fuzzy Systems 20 (1) (2018) 318–338.

[7] I. Eyoh, R. John, G. De Maere, Interval type-2 a-intuitionistic fuzzy logic for regression problems, IEEE Transactions on Fuzzy Systems 26 (4) (2018) 2396–2408.

[8] T.-C. Lu, Genetic-algorithm-based type reduction algorithm for interval type-2 fuzzy logic controllers, Engineering Applications of Artificial Intelligence 42 (2015) 36–44.

[9] O. Castillo, P. Melin, A review on interval type-2 fuzzy logic applications in intelligent control, Information Sciences 279 (2014) 615 – 631.

[10] K.-C. Kwak, A design of incremental granular model using context-based interval type-2 fuzzy c-means clustering algorithm, IEICE Transactions on Information and Systems E99D (1) (2016) 309–312.

[11] J. Xu, G. Feng, T. Zhao, X. Sun, M. Zhu, Remote sensing image classification based on semi-supervised adaptive interval type-2 fuzzy c-means algorithm, Computers and Geosciences 131 (2019) 132–143.

[12] J. Carlos Guzman, I. Miramontes, P. Melin, G. Prado-Arechiga, Optimal Genetic Design of Type-1 and Interval Type-2 Fuzzy Systems for Blood Pressure Level Classification, Axioms 8 (1).

[13] M. Antonelli, D. Bernardo, H. Hagras, F. Marcelloni, Multiobjective evolutionary optimization of type-2 fuzzy rule-based systems for financial data classification, IEEE Transactions on Fuzzy Systems 25 (2) (2017) 249–264.

[14] 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 Transactions on Fuzzy Systems 23 (4) (2015) 973–990.

[15] T. Pencheva, M. Angelova, K. Atanassov, Genetic algorithms quality assessment implementing intuitionistic Fuzzy logic, Vol. 3-4, 2015.

[16] T. Pencheva, M. Angelova, Intuitionistic fuzzy logic implementation to assess purposeful model parameters genesis, Studies in Computational Intelligence 657 (2017) 179–203.

[17] T. Kim, E. Sotirova, A. Shannon, V. Atanassova, K. Atanassov, L.-C. Jang, Interval Valued Intuitionistic Fuzzy Evaluations for Analysis of a Student's Knowledge in University e-Learning Courses, International Journal of Fuzzy Logic and Intelligent Systems 18 (3) (2018) 190–195.

[18] C. Zhang, Classification rule mining algorithm combining intuitionistic fuzzy rough sets and genetic algorithm, International Journal of Fuzzy Systems 22 (5) (2020) 1694–1715.

[19] J. Mendel, General type-2 fuzzy logic systems made simple: A tutorial, IEEE Transactions on Fuzzy Systems 22 (5) (2014) 1162–1182.

[20] J. Sanz, A. Fernndez, H. Bustince, F. Herrera, Improving the performance of fuzzy rule-based classification systems with interval-valued fuzzy sets and genetic amplitude tuning, Information Sciences 180 (19) (2010) 3674–3685.

[21] P. Melin, O. Castillo, A review on type-2 fuzzy logic applications in clustering, classification and pattern recognition, Applied Soft Computing 21 (2014) 568 – 577.

[22] J. H. Holland, Adaptation in Natural and Artificial Systems, University of Michigan Press, Ann Arbor, MI, 1975, second edition, 1992.

[23] 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 Transactions on Fuzzy Systems 19 (5) (2011) 857–872.

[24] R. Storn, K. Price, Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces, Journal of Global Optimization 11 (4) (1997) 341–359.

[25] G. Lucca, G. Dimuro, J. Fernandez, H. Bustince, B. Bedregal, J. Sanz, Improving the performance of fuzzy rule-based classification systems based on a nonaveraging generalization of CC-integrals named c_{F1F2}-integrals, IEEE Transactions on Fuzzy Systems 27 (1) (2019) 124–134.

[26] J. Hühn, E. Hüllermeier, FURIA: an algorithm for unordered fuzzy rule induction, Data Mining and Knowledge Discovery 19 (3) (2009) 293–319.

[27] J. Sanz, A. Fernández, 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 Transactions on Fuzzy Systems 21 (3) (2013) 399–411.

[28] J. Alcalá-Fdez, A. Fernández, J. Luengo, J. Derrac, S. García, L. Sánchez, F. Herrera, KEEL data-mining software tool: Data set repository, integration of algorithms and experimental analysis framework, Journal of Multiple-Valued Logic and Soft Computing 17 (2–3) (2011) 255–287.

[29] J. Demšar, Statistical comparisons of classifiers over multiple data sets, Journal of Machine Learning Research 7 (2006) 1–30.

[30] S. García, F. Herrera, An extension on "statistical comparisons of classifiers over multiple data sets" for all pairwise comparisons, Journal of Machine Learning Research 9 (2008) 2677–2694.

[31] S. García, A. Fernández, 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, Information Sciences 180 (10) (2010) 2044–2064.

[32] H. Bustince, E. Barrenechea, M. Pagola, J. Fernandez, Z. Xu, B. Bedregal, J. Montero, H. Hagras, F. Herrera, B. D. Baets, A historical account of types of fuzzy sets and their relationships, IEEE Transactions on Fuzzy Systems 24 (1) (2016) 179–194. `doi:10.1109/TFUZZ.2015.2451692`.

[33] H. Bustince, J. Fernandez, A. Kolesrov, R. Mesiar, Generation of linear orders for intervals by means of aggregation functions, Fuzzy Sets and Systems 220 (Supplement C) (2013) 69 – 77. doi:https://doi.org/10.1016/j.fss.2012.07.015.

[34] Z. S. Xu, R. R. Yager, Some geometric aggregation operators based on intuitionistic fuzzy sets, International Journal of General Systems 35 (4) (2006) 417–433.

[35] G. Beliakov, A. Pradera, T. Calvo, Aggregation Functions: A Guide for Practitioners. What is an aggregation function, Studies In Fuzziness and Soft Computing, Springer, San Mateo-California, 2007.

[36] H. Bustince, M. Galar, B. Bedregal, A. Kolesárová, R. Mesiar, A new approach to interval-valued Choquet integrals and the problem of ordering in interval-valued fuzzy set applications, Fuzzy Systems, IEEE Transactions on 21 (6) (2013) 1150–1162.

[37] G. Deschrijver, C. Cornelis, E. Kerre, On the representation of intuitionistic fuzzy t-norms and t-conorms, IEEE Transactions on Fuzzy Systems 12 (1) (2004) 45–61.

[38] G. Deschrijver, Generalized arithmetic operators and their relationship to t-norms in interval-valued fuzzy set theory, Fuzzy Sets and Systems 160 (21) (2009) 3080–3102.

[39] H. Ishibuchi, T. Nakashima, M. Nii, Classification and modeling with linguistic information granules: Advanced approaches to linguistic Data Mining, Springer-Verlag, Berlin, 2004.

[40] H. Ishibuchi, T. Nakashima, Effect of rule weights in fuzzy rule-based classification systems, IEEE Transactions on Fuzzy Systems 9 (4) (2001) 506–515.

[41] A. Jurio, M. Pagola, D. Paternain, C. Lopez-Molina, P. Melo-Pinto, Interval-valued restricted equivalence functions applied on clustering techniques, in: J. Carvalho, D. Kaymak, J. Sousa (Eds.), Proceedings of the joint 2009 international fuzzy systems association world congress and 2009 european society of fuzzy logic and technology conference, 2009, pp. 831–836, lisbon, Portugal, Jul. 20-24, 2009.

[42] O. Cordón, M. J. del Jesus, F. Herrera, A proposal on reasoning methods in fuzzy rule-based classification systems, International Journal of Approximate Reasoning 20 (1) (1999) 21–45.

[43] D. Wu, J. Mendel, Recommendations on designing practical interval type-2 fuzzy systems, Engineering Applications of Artificial Intelligence 85 (2019) 182–193.

[44] F. Olivas, L. Amador-Angulo, J. Perez, C. Caraveo, F. Valdez, O. Castillo, Comparative study of type-2 fuzzy particle swarm, bee colony and bat algorithms in optimization of fuzzy controllers, Algorithms 10 (3).

[45] T.-L. Le, T.-T. Huynh, C.-M. Lin, Self-evolving interval type-2 wavelet cerebellar model articulation control design for uncertain nonlinear systems using pso, International Journal of Fuzzy Systems 21 (8) (2019) 2524–2541.

[46] T. Zhao, Y. Xiang, S. Dian, R. Guo, S. Li, Hierarchical interval type-2 fuzzy path planning based on genetic optimization, Journal of Intelligent and Fuzzy Systems 39 (1) (2020) 937–948.

[47] J. Casillas, O. Cordon, F. Herrera, L. Magdalena, Interpretability issues in fuzzy modeling, Vol. 128, Springer, 2013.

[48] T. Miller, Explanation in artificial intelligence: Insights from the social sciences, Artificial Intelligence 267 (2019) 1–38.

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

[50] F. Herrera, M. Lozano, A. M. Sánchez, A taxonomy for the crossover operator for real-coded genetic algorithms: An experimental study, International Journal of Intelligent Systems 18 (3) (2003) 309–338.

[51] S. Z. C. Zhang, Association rule mining: Models and algorithms series, in: Lecture Notes in Computer Science, LNAI 2307, Berlin: Springer-Verlag, 2002.

[52] A. Segatori, A. Bechini, P. Ducange, F. Marcelloni, A distributed fuzzy associative classifier for big data, IEEE Transactions on Cybernetics 48 (9) (2018) 2656–2669.

[53] R. S. R. Agrawal, Fast algorithms for mining association rules, in: International Conference on Very Large Data Bases, Santiago de Chile, Chile, 1994, pp. 487–499.

[54] H. Ishibuchi, T. Yamamoto, Rule weight specification in fuzzy rule-based classification systems, IEEE Transactions on Fuzzy Systems 13 (4) (2005) 428–435.

[55] B. Kavek, N. Lavra, Apriori-sd: Adapting association rule learning to subgroup discovery, Applied Artificial Intelligence 20 (7) (2006) 543–583.

[56] J. Sanz, A. Fernández, H. Bustince, F. Herrera, IIVFDT: Ignorance functions based interval-valued fuzzy decision tree with genetic tuning, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 20 (supp02) (2012) 1–30.

[57] R. Alcalá, J. Alcalá-Fdez, M. J. Gacto, F. Herrera, Rule base reduction and genetic tuning of fuzzy systems based on the linguistic 3-tuples representation, Soft Computing 11 (5) (2007) 401–419.

[58] R.Storn, Differential evolution, a simple and efficient heuristic strategy for global optimization over continuous spaces, Journal of Global Optimization 11 (1997) 341–359.

[59] J. Alcalá-Fdez, A. Fernández, J. Luengo, J. Derrac, S. García, L. Sánchez, F. Herrera, KEEL data-mining software tool: Data set repository, integration of algorithms and experimental analysis framework, Journal of Multiple-Valued Logic and Soft Computing 17 (2–3) (2011) 255–287.

[60] T. Wong, P. Yeh, Reliable accuracy estimates from k-fold cross validation, IEEE Transactions on Knowledge and Data Engineering 32 (8) (2020) 1586–1594.

[61] D. Sheskin, Handbook of parametric and nonparametric statistical procedures, 2nd Edition, Chapman & Hall/CRC, 2006.

[62] J. L. Hodges, E. L. Lehmann, Ranks methods for combination of independent experiments in analysis of variance, Annals of Mathematical Statistics 33 (1962) 482–497.

[63] S. Holm, A simple sequentially rejective multiple test procedure, Scandinavian Journal of Statistics 6 (1979) 65–70.

[64] F. Wilcoxon, Individual comparisons by ranking methods, Biometrics 1 (1945) 80–83.

[65] T. Ho, M. Basu, Complexity measures of supervised classification problems, IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (3) (2002) 289–300.

[66] J. Luengo, F. Herrera, An automatic extraction method of the domains of competence for learning classifiers using data complexity measures, Knowledge and Information Systems 42 (1) (2015) 147–180.

[67] G. Lucca, J. Antonio Sanz, G. Dimuro, B. Bedregal, H. Bustince, R. Mesiar, C_F-integrals: A new family of pre-aggregation functions with application to fuzzy rule-based classification systems, Information Sciences 435 (2018) 94–110.

***Credit Author Statement**

J. Sanz has been involved in the following tasks: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data Curation, Writing - Original Draft, Visualization, Project administration

H. Bustince has been involved in the following tasks: Formal analysis, Investigation, Writing - Review & Editing, Visualization, Supervision, Funding acquisition