

Generating a Fuzzy Rule Based Classification System by genetic learning of granularity level using TOPSIS

Antonio E. de O. Araujo^a and Renato A. Krohling^b

^aGraduate Program in Computer Science - PPGI, Federal University of Espirito Santo - UFES, Vitoria, ES, Brazil, antonioeloy14@gmail.com

^bProduction Engineering Department, Federal University of Espirito Santo - UFES, Vitoria, ES, Brazil, krohling.renato@gmail.com

Abstract

Fuzzy Rule Based Classification Systems (FRBCSs) are widely used tools in classification problems. An important aspect in the design of a FRBCS is the number of fuzzy labels per variable (granularity level), which significantly influences the performance of the fuzzy system. Another relevant issue to be considered when generating a FRBCS is the accuracy-interpretability tradeoff, which can be addressed in the context of multi-objective optimization. Thus, in this work, we propose a new approach to design a FRBCS in which the accuracy and the interpretability (number of rules) of the FRBCS are considered objectives to be treated with the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS). We applied our method to several well-known standard classification datasets and the results show the feasibility of the proposed approach.

Keywords: Fuzzy Rule Based Classification Systems, Genetic Algorithm, Multi-objective optimization.

1 Introduction

Fuzzy Rule Based Classification Systems (FRBCSs) are an effective and widely used tools for pattern recognition and machine learning. These systems are able to achieve good accuracy results while providing interpretable models for the end user through the use of linguistic labels in the antecedents of their rules [16]. FRBCSs have been successfully applied in several areas including data mining [15, 8], network anomaly detection [11], medical diagnosis [27], business process [26] and classification of grape berries [21].

The generation of the Knowledge Base (KB) of a FRBCS is usually done using the knowledge of a specialist, or through an automatic learning process that aims to derive the KB from the samples. Many approaches have been proposed to automatically learn the KB of a FRBCS. Most of them consist of obtaining a Rule Base (RB) from a predefined Data Base (DB), in which all variables have the same number of linguistic terms, typically an odd number between 3 and 7. However, the choice of the number of linguistic terms for each fuzzy variable (granularity level) has a significant influence on the performance of a fuzzy system. Therefore, some methods for the automatic generation of the KB in fuzzy classification [31] and fuzzy modeling [5] include granularity level learning.

Another important aspect in the design of a FRBCS is related to accuracy and interpretability. During the generation of a FRBCS, we seek to obtain a linguistic model with high accuracy as well as high interpretability. However, these two objectives are considered to be conflicting. Improvement in any of these two objectives can cause a decrement in the another one [7]. Therefore, the design of a FRBCS can be seen as a tradeoff between the accuracy and interpretability of the model. In some cases, one may prefer a complex linguistic model with high accuracy, in others desire a simpler one with high interpretability [16]. Thus, the generation of a FRBCS is also treated in the context of multi-objective optimization. Multi-objective genetic algorithms have already been used in obtaining FRBCSs with different accuracy and interpretability [3], [13].

In this work, we present a new approach to design a FRBCS in the context of multi-objective optimization, in which the accuracy over the training dataset and the number of rules of the FRBCS are considered objectives to be treated with the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS).

The remainder of this paper is organized as follows. First, in Section II, we present a brief background on

FRBCS. Section III consists of a detailed description of the proposed approach. Section IV presents the experimental results. Finally, in Section V, some conclusions and future works are pointed out.

2 Background

2.1 Fuzzy Rule Based Classification Systems

A classification problem is usually defined by a set of m training samples $x_p = (x_{p1}, \dots, x_{pn}, C_p)$, $p = 1, 2, \dots, m$, and M classes, where x_{pi} is the value of the i -th attribute and C_p is the class label of the p -th training sample [6].

A FRBCS is composed of two parts: the Fuzzy Reasoning Method (FRM) and the Knowledge Base (KB). The KB is divided into two main parts: the Data Base (DB), formed by the membership functions of the fuzzy sets associated to each linguistic variable, and the Rule Base (RB), composed of a set of fuzzy rules in the following format:

$$\text{Rule } R_j : \text{If } x_1 \text{ is } A_{j1} \text{ and } \dots \text{ and } x_n \text{ is } A_{jn} \quad (1) \\ \text{then Class} = C_j \text{ with } RW_j$$

where R_j is the label of the j -th rule, $x = (x_1, \dots, x_n)$ is a n -dimensional vector representing a given sample, A_{ji} is an antecedent fuzzy set for the i -th attribute, C_j is the class label and RW_j is the rule weight.

The weight of the fuzzy rule significantly influences the performance of a FRBCS [17]. Several heuristic methods were proposed to determine it. In this work, we use the method known as Penalized Certainty Factor (PCF) [19], defined as follows:

$$RW_j = PCF_j = \frac{\sum_{x_p \in C_j} \mu_{A_j}(x_p) - \sum_{x_p \notin C_j} \mu_{A_j}(x_p)}{\sum_{p=1}^m \mu_{A_j}(x_p)} \quad (2)$$

where $\mu_{A_j}(x_p)$ is the membership degree of the sample x_p of the training set in relation to the antecedents of the rule j , and C_j is the label of the class present in the consequent of the rule j .

In order to classify a new sample, we adopted the single winner method [18]. Considering that the RB is formed by a set of S fuzzy rules in the format specified in (1), we first determine the winning rule R_ω for a given sample $x_p = (x_{p1}, \dots, x_{pn})$ according to the following equation:

$$\mu_\omega(x_p) \cdot RW_\omega = \max \{ \mu_j(x_p) \cdot RW_j; j = 1 \dots S \} \quad (3)$$

Then, we associate to the new sample the class present in the consequent of the winning rule R_ω , in this case, the class label C_j .

2.2 Chi et al.'s algorithm

The approach for generating the KB of the FRBCS proposed in this paper uses the Chi et al.'s algorithm [4] to derive the RB of the classifier. This method is an extension of the Wang and Mendel algorithm [32] to classification problems. The main steps of this algorithm are described in the following [6]:

1. *Definition of the fuzzy partitions:* The fuzzy partitions are defined according to the domain associated with each attribute of the dataset. In this work, we use uniform fuzzy partitions with triangular membership functions.
2. *Generation of a fuzzy rule associated with each sample $x_p = (x_{p1}, \dots, x_{pn}, C_p)$:*
 - (a) Calculate the membership degree μ_{x_p} of the sample with respect to the fuzzy partitions associated with each input attribute (in this work, we chose the product T-norm).
 - (b) Assign the sample x_p to the fuzzy region with the greatest membership degree.
 - (c) Construct a rule for the sample, whose antecedent consists of the fuzzy region established previously and consequent is the class label C_j of the sample.
 - (d) Calculate the rule weight RW_j .

During the learning process, it is possible to generate rules with the same antecedent. If they have the same consequent, we simply remove one of the duplicated rules. Otherwise, we keep in the RB only the rule with the greatest weight.

3 Proposed Approach

In this section we describe in detail the proposed approach to generating a FRBCS by genetic learning of granularity level using TOPSIS. We use the CHC algorithm [10] for the granularity level learning, the Chi et al.'s method [4] for the RB derivation and the single winner method [18] as the FRM. As our main purpose is to illustrate the effectiveness of the proposed approach, we decided to employ a quick RB derivation method and a classical FRM. However, a more accurate fuzzy rules generation algorithm such as FARCHD [1] may be used, as well as a FRM that adopts the Choquet Integral or its generalizations [22], [23], [20].

Our method is based on CHC-GL algorithm, that was performed in [31] for multi-class problems. The main

difference of our approach is the use of TOPSIS with multi-criteria information in the evaluation of the individuals of the evolutionary algorithm. Therefore, we denote our proposal as CHC-GL-TOPSIS (CHC for Granularity Learning using TOPSIS). Next, we describe the main concepts related to TOPSIS and also the structure of the Genetic Algorithm (GA) [12] used in the proposed approach for the genetic learning of the KB.

3.1 TOPSIS

TOPSIS is a technique proposed in 1981 by Hwang and Yoon [14] that allows to determine the best alternative in a problem that involves several decision criteria, called multi-criteria decision making (MCDM) problem. The best alternative is the one whose distance to the positive ideal solution (PIS) is the smallest and the distance to the negative ideal solution (NIS) is the largest. The PIS is the solution that contains the best values for each criterion considering all possible alternatives, while the NIS is the solution that has the worst values [28].

To find the best alternative of a MCDM problem using TOPSIS, we must first define a decision matrix D with m alternatives and n criteria that can be described by

$$D = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix} \quad (4)$$

where x_{ij} is the rating of the i -th alternative under the j -th criterion. The criteria are weighted by a vector of weights W described by

$$W = (w_1, w_2, \dots, w_j), w_j \geq 0 \quad (5)$$

where w_j is the weight of the j -th criterion, satisfying $\sum_{j=1}^n w_j = 1$. A criterion whose value is to be maximized is considered as *benefit*. In contrast, when the value of the criterion should be minimized, it is considered as *cost*.

Since the criteria have different origins, it is necessary to normalize the decision matrix D so that it is possible to compare the values relative to different criteria considered [9]. In this work, the decision matrix D is normalized for each criterion according to

$$p_{ij} = \frac{x_{ij} - \min_i(x_{ij})}{\max_i(x_{ij}) - \min_i(x_{ij})}, i = 1, \dots, m, j = 1, \dots, n \quad (6)$$

Thus, the normalized decision matrix A_n represents the relative performance of each alternative and can be described by $A_n = (p_{ij})_{m \times n}, i = 1, \dots, m, j = 1, \dots, n$

The TOPSIS algorithm begins with the calculation of the positive ideal solution A^+ and negative ideal solution A^- .

$$A^+ = (p_1^+, p_2^+, \dots, p_n^+) \quad (7)$$

$$A^- = (p_1^-, p_2^-, \dots, p_n^-) \quad (8)$$

where

$$p_j^+ = (\max_i p_{ij}, j \in J_1; \min_i p_{ij}, j \in J_2) \quad (9)$$

$$p_j^- = (\min_i p_{ij}, j \in J_1; \max_i p_{ij}, j \in J_2) \quad (10)$$

J_1 and J_2 represent the *benefit* and *cost* criteria, respectively.

Then, the Euclidean distance between A_i and A^+ , and between A_i and A^- is calculated, respectively, according to the following equations:

$$d_i^+ = \sqrt{\sum_{j=1}^n w_j (p_j^+ - p_{ij})^2} \quad (11)$$

$$d_i^- = \sqrt{\sum_{j=1}^n w_j (p_j^- - p_{ij})^2} \quad (12)$$

Next, the relative closeness ξ_i between each i -th alternative and A^+ is calculated according to

$$\xi_i = \frac{d_i^-}{d_i^+ + d_i^-} \quad (13)$$

Finally, the best alternative is chosen, which is the one closest to PIS, that is, with the greatest value of ξ_i .

3.2 Genetic Learning of the KB

The genetic learning of the KB is carried out with the use of the CHC algorithm [10], a binary-coded elitist genetic algorithm that combines an elitist selection strategy with a heterogeneous uniform crossover, promoting a high diversity into the population [30]. This algorithm has already been successfully used in multi-class problems [31] and also in big data classification [25]. In the following paragraphs we describe in detail the main properties of the proposed CHC algorithm.

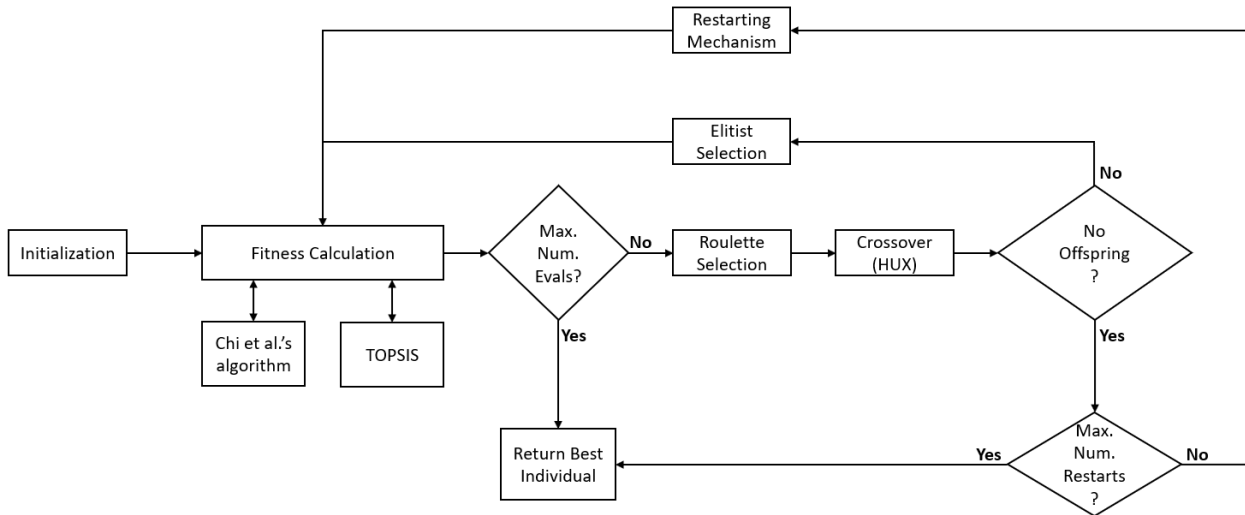


Figure 1: Flowchart of the proposed CHC algorithm.

The general flowchart of the proposed CHC scheme is presented in Figure 1.

The individuals of the GA codify the granularity level of each attribute. Therefore, the length of the individual is equal to the number of attributes of the dataset. Each gene represents the amount of fuzzy labels to be used in each input variable partition. In this paper, the number of labels is restricted to the values that belong to the set $S = \{2, \dots, 7\}$. Therefore, an individual I can be represented as

$$I = (g_1, g_2, \dots, g_N) \quad (14)$$

where g_i is the value that represents the granularity level of the linguistic variable i .

The initial population is formed in two stages. In the first one, six individuals with the same granularity level for all linguistic variables are created. Each individual corresponds to one of the six possible values for the number of labels defined in $S = \{2, \dots, 7\}$. The remaining individuals of the population are randomly generated according to the possible values for the granularity level.

The calculation of the fitness value of the individuals is performed through the following steps:

- For each individual, define the whole DB using the granularity level encoded in the individual. The domain of each attribute is represented by a uniform fuzzy partition with triangular membership functions, according to the respective number of labels specified in g_i .
- Generate the RB for each individual using the Chi

et al.'s algorithm.

- From the newly formed KB, generate a FRBCS associated with each individual.
- Calculate the accuracy over the training dataset of each FRBCS generated.
- Calculate the fitness value for each individual. In our work, we consider the problem of evaluating each individual as a MCDM problem, in which each individual represents an alternative, the *Accuracy over the Training Dataset* of the FRBCS associated to the individual is considered a *benefit* criterion, and the *Number of Rules* of the RB is taken as a *cost* criterion. Using TOPSIS, the fitness value of an individual is the relative closeness value ξ . The best alternative, i.e., the best individual, is the one with the greatest value of ξ .

The mechanism of selection used in the algorithm is called *Elitist Selection*. If M is the size of the population, in each generation of the algorithm a new population is formed from the selection of the M best individuals considering the current population and its offspring.

The *Incest prevention* property aims to maintain the genetic diversity of the population. Two parents can be crossed only if the Hamming distance between them divided by two is greater than a predetermined threshold T , which is usually initialized to $L/4$, being L the length of the individual. If no individual from offspring is selected for the next generation, the threshold value is decreased by one.

Whenever the threshold value T is equal to 0, a *Restarting process* occurs, in which the current popu-

lation is reinitialized. The best individual is preserved and the remaining individuals of the new population are created by randomly changing a percentage (defined by the divergence rate dr) of the best individual's genes [29].

Finally, the crossover operator applied in the algorithm proposed is the *Half Uniform Crossover (HUX)*. In this way, exactly half of the different genes between two parents are exchanged. These genes are randomly selected and cause the offspring generated to be as different as possible from their parents, which increases the diversity of the population and decreases the likelihood of premature convergence of the GA. The pseudo-code of the CHC algorithm is described in Algorithm 1.

Algorithm 1 CHC algorithm

Input: $M, L, dr, maxEval, maxRestart$

Output: KB

```

1:  $numEval = 0$ 
2:  $numRestart = 0$ 
3:  $threshold = L/4$ 
4:  $P = initPopulation(M)$ 
5:  $F = evalPopulation(P)$ 
6: for ( $numEval < maxEval$ ) and ( $numRestart < maxRestart$ ) do
7:    $P' = select(P)$ 
8:    $P'' = crossover(P')$ 
9:   if ( $isEmpty(P'') == 1$ ) then
10:     $threshold = threshold - 1$ 
11:    if ( $threshold < 0$ ) then
12:       $P = restart(M, dr)$ 
13:       $F = evalPopulation(P)$ 
14:       $numRestart = numRestart + 1$ 
15:       $threshold = L/4$ 
16:    end if
17:  else
18:     $P = elitistSelection(P', P'')$ 
19:     $F = evalPopulation(P)$ 
20:     $numEval = numEval + 1$ 
21:  end if
22: end for
23:  $KB = bestIndividual(P)$ 
24: return  $KB$ 

```

4 Experimental Results

The main objective of this experimental study is to show that the genetic learning of the granularity level using TOPSIS is able to increase the accuracy over the test dataset and reduce the number of rules of a FRBCS, improving its prediction ability and interpretability.

To show the feasibility of the proposed approach, we will compare the performance of the FRBCS generated by the CHC-GL-TOPSIS method with three other classifiers, whose RB is derived by the Chi et al. algorithm and the DB is formed by uniform fuzzy partitions and triangular membership functions, with 3,

5 and 7 labels per variable, denoted Chi-G3, Chi-G5 and Chi-G7, respectively.

4.1 Datasets

The experiments were performed with ten standard classification datasets from KEEL dataset repository [2]. The main characteristics of each dataset are presented in Table 1. In order to avoid covariate-shift-related problems, each dataset was divided into 5 different partitions according to the Distribution Optimally Balanced Cross Validation (DOB-SCV) method [24].

Dataset	Samples	Attributes	Classes
Appendicitis	106	7	2
Haberman	306	3	2
Balance	625	4	3
Hayes-Roth	160	4	3
Iris	150	4	3
Newthyroid	215	5	3
Post-Operative	87	8	3
Tae	151	5	3
Wine	178	13	3
Zoo	101	16	7

Table 1: Characteristics of datasets.

4.2 Experimental setup

The parameters of the CHC algorithm used in the proposed approach are listed in Table 2.

Parameter	Value
Maximum number of evaluations	1000
Population size	50
Divergence rate	0.35

Table 2: CHC parameters.

A sensitivity analysis was performed to determine the weights of the evaluation criteria used in TOPSIS [9]. The analysis consisted in the execution of the algorithm using the *Accuracy over the Training Dataset* and the *Number of Rules* of the FRBCS as TOPSIS criteria for the 10 datasets and for vectors of weight in intervals of 0.05. The results of accuracy obtained are presented in Figure 2. From these results we chose the weight vector $W = (0.60, 0.40)$.

4.3 Results Analysis

In Table 3, the mean accuracy over the test dataset and the average number of rules are presented for each FRBCS. The best values are shown in bold.

As it can be observed, the FRBCS generated by the CHC-GL-TOPSIS method obtained a higher accu-

Dataset	Chi-G3		Chi-G5		Chi-G7		CHC-GL-TOPSIS	
	Accuracy (%)	Rules	Accuracy (%)	Rules	Accuracy (%)	Rules	Accuracy (%)	Rules
Appendicitis	87.79	28.4	84.94	56.8	85.93	73,4	87.75	44.2
Haberman	71.91	11.4	71.58	41.8	73.54	78,0	75.51	16.4
Balance	90.40	68.4	46.08	500.0	46.08	500,0	90.72	67.8
Hayes-Roth	55.00	41.4	58.83	70.0	60.04	69,6	64.97	58.4
Iris	86.67	13.4	95.33	42.6	94.00	62,6	96.67	10.4
Newthyroid	90.70	18.8	93.49	44.0	92.09	62,6	90.70	9.6
Post-Operative	67.90	55.2	66.97	55.8	66.72	55,8	64.62	49.6
Tae	58.31	23.8	59.60	56.4	54.19	58,4	53.50	46.2
Wine	91.58	114.4	77.00	141.8	39.39	142,4	89.93	123.2
Zoo	73.31	49.4	72.66	50.4	71.03	51,0	75.20	48.0

Table 3: Accuracy over the test dataset and number of rules of the FRBCSs.

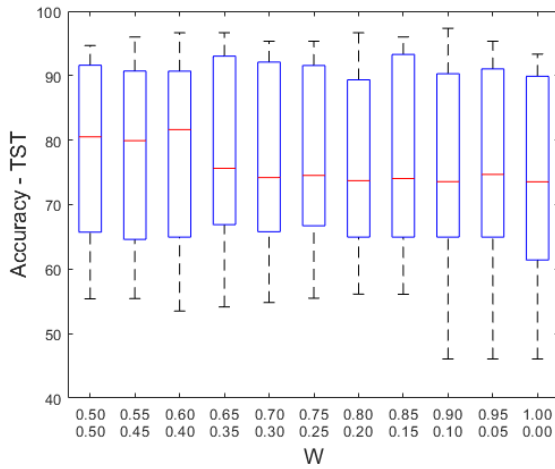


Figure 2: Weight vector sensitivity analysis.

racy over the test dataset in 5 out of 10 datasets used (Haberman, Balance, Hayes-Roth, Iris and Zoo). Considering the FRBCS generated by methods without granularity level learning, Chi-G3 obtained the greatest accuracy in 3 datasets (Appendicitis, Post-Operative and Wine), while Chi-G5 presented better prediction ability in 2 datasets (Newthyroid and Tae). These results show the importance of the granularity level in the performance of the fuzzy classifier, and also suggest that the proposed approach can generate FRBCSs that present good accuracy results in classification of datasets with few attributes.

Regarding the average number of rules, we can note that the CHC-GL-TOPSIS method was able to generate FRBCS with the smallest RB in 5 out of 10 datasets used in the experiments (Balance, Iris, Newthyroid, Post-Operative and Zoo). A similar result was obtained with the FRBCS generated by the Chi-G3 method, which presented the smallest RB in the remaining 5 datasets. As expected, for the FRBCS generated by methods without granularity level

learning (Chi-G3, Chi-G5 and Chi-G7), the number of rules of the RB increases as the number of labels per linguistic variable increases.

It's important to note that, in 3 out of 10 datasets (Balance, Iris and Zoo), the FRBCS generated by the CHC-GL-TOPSIS method achieved the best accuracy over the test dataset and also presented the RB with the smallest number of rules, increasing the prediction ability and the interpretability of the classifier in relation to FRBCS generated by methods without granularity level learning.

5 Conclusion

In this paper, a method for generating a FRBCS by genetic learning of granularity level using TOPSIS was proposed. This method, which we call CHC-GL-TOPSIS, was applied to several standard classification datasets and compared to classical methods without granularity level learning. The experimental results showed that the CHC-GL-TOPSIS method is able to increase the accuracy over the test dataset and reduce the number of rules of a FRBCS, indicating the feasibility of the proposed approach. We must remark that we used a simple algorithm for RB derivation, but more accurate methods can be employed. Due to computational time required for training, we conducted the experimental study just with standard classification datasets. In the future, we will work to employ a more accurate algorithm for RB generation and to carry out the experiments with large datasets.

Acknowledgement

R.A. Krohling would like to thank the financial support of the Brazilian agency CNPq and the local Agency of the state of Espirito Santo FAPES under grant nr. 309161/2015-0 and nr. 039/2016, respectively.

References

- [1] J. Alcalá-Fdez, R. Alcalá, 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.
- [2] 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 Multi-Valued Logic and Soft Computing* 17 (2010) 255–287.
- [3] E. H. Cárdenas, H. A. Camargo, Y. J. Túpac, Imbalanced datasets in the generation of fuzzy classification systems - an investigation using a multi-objective evolutionary algorithm based on decomposition, in: 2016 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Vancouver, BC, 2016, pp. 1445–1452.
- [4] Z. Chi, H. Yan, T. Pham, *Fuzzy Algorithms: With Applications to Image Processing and Pattern Recognition*, World Scientific, 1996.
- [5] O. Cordon, F. Herrera, P. Villar, Generating the knowledge base of a fuzzy rule-based system by the genetic learning of the data base, *IEEE Transactions on Fuzzy Systems* 9 (4) (2001) 667–674.
- [6] S. del Río, V. Lopez, J. M. Benítez, F. Herrera, A mapreduce approach to address big data classification problems based on the fusion of linguistic fuzzy rules, *International Journal of Computational Intelligence Systems* 8.
- [7] P. K. Dwivedi, S. P. Tripathi, A survey on the design of fuzzy classifiers using multi-objective evolutionary algorithms, *International Journal of Science and Research (IJSR)* 4 (2015) 1103–1107.
- [8] M. Elkan, M. Galar, J. Sanz, H. Bustince, Chibd: A fuzzy rule-based classification system for big data classification problems, *Fuzzy Sets and Systems* 348 (2018) 75–101.
- [9] J. G. M. Esgario, R. A. Krohling, Clustering with minimum spanning tree using topsis with multi-criteria information, in: 2018 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Rio de Janeiro, 2018, pp. 1–7.
- [10] L. J. Eshelman, The chc adaptive search algorithm: How to have safe search when engaging in nontraditional genetic recombination, in: G. J. E. Rawlins (Ed.), *Foundations of Genetic Algorithms*, Vol. 1, Elsevier, 1991, pp. 265–283.
- [11] S. Ghosh, A. Pal, A. Nag, S. Sadhu, R. Pati, Network anomaly detection using a fuzzy rule-based classifier, in: *Computer Communication and Electrical Technology*, 2017, pp. 61–65.
- [12] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, 1989.
- [13] C. E. Hinojosa, H. A. Camargo, V. Y. J. Túpac, Learning fuzzy classification rules from imbalanced datasets using multi-objective evolutionary algorithm, in: 2015 Latin America Congress on Computational Intelligence (LA-CCI), Curitiba, 2015, pp. 1–6.
- [14] C. L. Hwang, K. P. Yoon, *Multiple attribute decision making: methods and applications*, Springer-Verlag, Berlin, 1981.
- [15] L. Iníguez, M. Galar, A. Fernández, Improving fuzzy rule based classification systems in big data via support-based filtering, in: 2018 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Rio de Janeiro, 2018, pp. 1–8.
- [16] H. Ishibuchi, *Classification and Modeling with Linguistic Information Granules: Advanced Approaches to Linguistic Data Mining*, Springer-Verlag, Berlin, Heidelberg, 2009.
- [17] H. Ishibuchi, T. Nakashima, Effect of rule weights in fuzzy rule-based classification systems, *IEEE Transactions on Fuzzy Systems* 9 (4) (2001) 506–515.
- [18] H. Ishibuchi, T. Nakashima, T. Morisawa, Voting in fuzzy rule-based systems for pattern classification problems, *Fuzzy Sets and Systems* 103 (2) (1999) 223–238.
- [19] H. Ishibuchi, T. Yamamoto, Rule weight specification in fuzzy rule-based classification systems, *IEEE Transactions on Fuzzy Systems* 13 (4) (2005) 428–435.
- [20] G. Lucca, G. P. Dimuro, J. Fernández, H. Bustince, B. Bedregal, J. A. Sanz, Improving the performance of fuzzy rule-based classification systems based on a nonaveraging generalization of CC -integrals named $C_{F_1 F_2}$ -integrals, *IEEE Transactions on Fuzzy Systems* 27 (1) (2019) 124–134.
- [21] G. Lucca, J. A. Sanz, H. Bustince, G. P. Dimuro, V. Gomes, R. C. C. Madureira, P. Melo-Pinto, Applying aggregation and pre-aggregation functions in the classification of grape berries, in: 2018 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Rio de Janeiro, 2018, pp. 1–6.

- [22] G. Lucca, J. A. Sanz, G. P. Dimuro, B. Bedregal, M. J. Asiain, M. Elkano, H. Bustince, CC-integrals: Choquet-like copula-based aggregation functions and its application in fuzzy rule-based classification systems, *Knowledge-Based Systems* 119 (2017) 32–43.
- [23] G. Lucca, J. A. Sanz, G. P. 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.
- [24] J. G. Moreno-Torres, J. A. Saez, F. Herrera, Study on the impact of partition-induced dataset shift on k -fold cross-validation, *IEEE Transactions on Neural Networks and Learning Systems* 23 (8) (2012) 1304–1312.
- [25] D. Peralta, S. Río, S. Ramírez, I. Triguero, J. M. Benítez, F. Herrera, Evolutionary feature selection for big data classification: A mapreduce approach, *Mathematical Problems in Engineering* 2015.
- [26] L. N. Safitri, R. Sarno, G. I. Budiawati, Improving business process by evaluating enterprise sustainability indicators using fuzzy rule based classification, in: *2018 International Seminar on Application for Technology of Information and Communication*, Semarang, 2018, pp. 55–60.
- [27] J. A. Sanz, M. Galar, A. Jurio, A. Brugos, M. Pagola, H. Bustince, Medical diagnosis of cardiovascular diseases using an interval-valued fuzzy rule-based classification system, *Applied Soft Computing* 20.
- [28] E. R. Schneider, R. A. Krohling, A hybrid approach using topsis, differential evolution, and tabu search to find multiple solutions of constrained non-linear integer optimization problems, *Knowledge-Based Systems* 62 (2014) 47–56.
- [29] A. Simões, E. Costa, Memory-based chc algorithms for the dynamic traveling salesman problem, in: *Genetic and Evolutionary Computation Conference, GECCO'11*, New York, 2011, pp. 1037–1044.
- [30] I. Triguero, M. Galar, H. Bustince, F. Herrera, A first attempt on global evolutionary undersampling for imbalanced big data, in: *2017 IEEE Congress on Evolutionary Computation (CEC)*, San Sebastian, 2017, pp. 2054–2061.
- [31] P. Villar, A. Fernández, R. Montes, A. M. Sánchez, F. Herrera, Improving the ovo performance in fuzzy rule-based classification systems by the genetic learning of the granularity level, in: *2015 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, Istanbul, 2015, pp. 1–7.
- [32] L. X. Wang, J. M. Mendel, Generating fuzzy rules by learning from examples, *IEEE Transactions on Systems, Man, and Cybernetics* 22 (6) (1992) 1414–1427.