Fuzzy Classification Rules with FRvarPSO Using Various Methods for Obtaining Fuzzy Sets

Patricia Jimbo Santana
Accounting and Auditing Branch, School of Administration Science, Central University of Ecuador, Quito, Ecuador
Email: prjimbo@uce.edu.ec

Laura Lanzarini
Computer Science Research Institute LIDI (III-LIDI), School of Computer Science, National University of La Plata, La Plata, Buenos Aires, Argentina
Email: laural@lidi.info.unlp.edu.ar

Aurelio F. Bariviera
Department of Business, Rovira i Virgili University, Avenida de la Universitat, 1 Reus, Spain
Email: aurelio.fernandez@urv.cat

***Abstract*—Having strategies capable of automatically generating classification rules is highly useful in any decision-making process. In this article, we propose a method that can operate on nominal and numeric attributes to obtain fuzzy classification rules by combining a competitive neural network with an optimization technique based on variable population particle swarms. The fitness function that controls swarm movement uses a voting criterion that weights, in a fuzzy manner, numeric attribute participation. The efficiency and efficacy of this method are strongly conditioned by how membership functions to each of the fuzzy sets are established. In previous works, this was done by partitioning the range of each numeric attribute at equal-length intervals, centering a triangular function with appropriate overlap in each of them. In this case, an improvement to the fuzzy set generation process is proposed using the Fuzzy C-Means methods. The results obtained were compared to those yielded by the previous version using 11 databases from the UCI repository and three databases from the Ecuadorian financial system – one from a credit and savings cooperative and two from banks that grant productive and non-productive credits as well as microcredits. The results obtained were satisfactory. At the end of the article, our conclusions are discussed and future research lines are suggested.**

***Index Terms*—FRvarPSO (Fuzzy Rules variable Particle Swarm Optimization), fuzzy rules, classification rules, fuzzy C-means, data mining**

I. Introduction

Data mining allows exploring large volumes of data with the purpose of extracting existing patterns or relations. The technique to be used in each case depends on the type of problem to be solved. When classifying available examples or cases, classification rules are usually one of the most widely used models. A classification rule is a conditional expression following an IF-THEN structure, where the antecedent is a conjunction of conditions that allow identifying the reasons why an example is considered to belong to a given class [1]. In this article, linguistic variables have been used to express numeric attributes. This fuzzy representation increases rule coverage and allows applying it even if the example being classified does not fully meet the conditions included in the antecedent. Additionally, the lack of hard boundaries to analyze numeric variables helps understand the rule, resulting in a more intuitive use. The following is an example of a classification rule:

IF “ability to pay = low” AND “delinquency = medium” AND “Requested amount = medium” THEN credit = “N”

In this article, we propose generating the set of classification rules using the FRvarPSO (Fuzzy Rules variable Particle Swarm Optimization) method. This method was defined in [2] and can operate on both nominal and numeric attributes to obtain fuzzy classification rules. Its operation is based on combining a competitive neural network with a variable population particle swarm-based optimization technique. The fitness function that controls swarm movement uses a voting criterion that weights, in a fuzzy manner, numeric attribute participation. As a result of applying FRvarPSO, a set of fuzzy classification rules with low cardinality is obtained. These rules are typically easy to understand and are pretty accurate.

The efficiency and efficacy of this method are strongly conditioned by how membership functions to each of the fuzzy sets are established. In previous works, this was done by partitioning the range of each numeric attribute at equal-length intervals, centering a triangular function with appropriate overlap in each of them [3].

However, according to [4], to solve the issue of extracting knowledge, the linguistic labels and their fuzzy sets should be defined in collaboration with the expert, obtaining as a result a reduced granularity in attribute domains. Since consulting with an expert is not always

Manuscript received April 29, 2020; revised September 27, 2020.

doi: 10.12720/jait.11.4.233-240

possible, we are proposing a method to improve the process for obtaining fuzzy sets using Fuzzy C-Means and then applying FRvarPSO to get the fuzzy rules.

The remaining sections of this article are organized as follows: Section II describes some related articles, Section III details the method proposed, Section IV presents the results obtained, and Section V summarizes our conclusions and proposes some future lines of work.

II. Related Work

In the literature, there are several works that combine a number of techniques for obtaining fuzzy rules. The difference between those articles and our work lies in the regions where the examples representing the fuzzy sets are found, the structure of the membership functions, and the techniques used to obtain the fuzzy rules.

Among these articles, the authors in [5] use Fuzzy C-means (FCM) to establish fuzzy rules to estimate temperature, and combine it with learning functions to obtain good results. Similarly, the authors in [6] work with variations of FCM and consider it to be one of the most accurate and efficient algorithms to be used in clustering problems. In [7], a probabilistic neural network neuronal is combined and pattern neurons are selected based on the centroids obtained, also using the highest value of the membership functions for the fuzzy sets to activate pattern neurons. In [8], the FCM algorithm is used to establish the fuzzy sets in combination with genetic algorithms to obtain the rules, and then applying techniques for rule reduction, achieving good accuracy. In [9], the authors focus on fuzzy rule weights to consider that they belong to the rule system, rather than establishing fuzzy sets. In [10], simple or equitable fuzzy partitions are used, combined with hybrid methods such as the Apriori algorithm and genetic algorithms, testing only on two databases from the UCI repository. Later on, a different type of research work was carried out, such as the one discussed in [11], which use neuro-fuzzy systems where the membership function is obtained through a bell-shaped distribution. This method uses a fuzzification matrix where input patterns are associated to a degree of membership to different classes. Based on the degree of membership, a pattern would be attributed to a specific category or class.

In [12], the multiple-input, single-output, non-linear fuzzy systems proposed by Takagy Sugeno [13], are used. Fuzzy sets are built by clustering, using the Fuzzy C-means algorithm. In other articles, authors highlight the importance of fuzzy clustering for fuzzy modeling [14]-[17], and suggest that new fuzzy models should be created using fuzzy classification rules, which is the foundation for our article.

Other research works that have yielded excellent results are based on image clustering, obtaining robust models [18], [19]. In [20], the authors mention how hard it is, for traditional, non-supervised classification algorithms, to create an accurate classification model, and how using clustering with FCM allows effectively managing data defuzzification. The authors in [21] suggest that FCM should be used when the dataset at hand is hard to cluster, and that combining it with other algorithms such as K-means can improve model efficiency.

In several types of applications, such as biology-related ones [22], hybrid clustering methods are also used, including FCM, which significantly increases accuracy compared to other methods that do not use FCM. Recent works such as the one discussed in [23] corroborate the fact that, even though FCM is used in fuzzy rule generation, especially to build the fuzzy sets formed by rule conditions, this does not contribute directly to system modeling, so variations are required.

The reason why Fuzzy C-means is used in our work to obtain the fuzzy sets required to express the conditions involving numeric attributes is that an expert is not always available to collaborate. This technique, combined with FRvarPSO, not only allows obtaining fuzzy classification rules, but it also helps achieve a balance between rule accuracy and simplicity.

III. Methodology

A. Fuzzy C-Means (FCM)

Among data mining techniques, clustering algorithms are an essential tool when solving descriptive problems. They help gather examples with common characteristics to generate a model based on groups of similar elements based on a given metric. By analyzing and observing these groups, patterns or relations can be inferred and then used to describe them.

In this article, the Fuzzy C-means (FCM) algorithm [24] will be used as clustering technique. As a result, a centroid-based model will be obtained, and functions will be used to measure the degree of membership of each example to each group. Thus, unlike partitive clustering methods, a single example may belong to more than one group or cluster, and will carry an associated degree of membership accordingly. Thus, the examples will partially belong to each of the groups, and fuzzy partitions will be created.

Each cluster is considered as a fuzzy set. One of the input parameters for the algorithm is the number of clusters C. Once it is applied, the centroids and the corresponding membership functions are defined. The following considerations should be taken into account:

- The sum of the degrees of membership of the ith example to the different fuzzy sets is 1.
- The degree of membership of the ith example in cluster, called μ_ij, must be between 0 and 1.

The goal of the FCM algorithm is minimizing the following target function:

$$J=\sum_{i=1}^{N}\sum_{j=1}^{C}\mu_{ij}\cdot\left\|X_i - C_i\right\|^2 \tag{1}$$

where N is the number of examples, C is the number of clusters to build, C_j is the centroid of cluster j, μ_{ij} is the degree of membership for the i^{th} example X_i. In cluster j and euclidean norm $X_i - C_j$ responsible for measuring how close example X_i is to cluster j.

The degree of membership μ_{ij} of example X_i, to cluster j, is calculated as follows:

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{c}\left[\frac{\left\|X_i - C_j\right\|}{\left\|X_j - C_k\right\|}\right]^{\frac{2}{m-1}}} \tag{2}$$

where m is considered to be the fuzzification coefficient. This value indicates how many clusters may be overlapped. As m increases, so will cluster overlap.

Centroid C_j is obtained as follows:

$$C_j = \frac{\sum_{i=1}^{N}\mu_{ij}^{m}.X_j}{\sum_{i=1}^{N}\mu_{ij}^{m}} \tag{3}$$

μ_{ij} is in the interval [0, 1], so that:

$$\sum_{j=1}^{C}\mu_{ij} = 1 \tag{4}$$

The Fuzzy C-means algorithm is used to generate the membership function of the fuzzy/linguistic variable to each of the fuzzy sets, without carrying out an equitable distribution of the fuzzy sets, or when there is no expert available to consult. For instance, the fuzzy variable “delinquency” is measured in “days”, a discourse universe (minimum value, maximum value), 3 clusters, with linguistic terms (low, medium, high), each of these terms being determined by a fuzzy set with a membership function.

B. Fuzzy Rules Variable Particle Swarm Optimization (FRvarPSO)

The method (FRvarPSO) was defined in [2], [25], [26], and its goal is obtaining a set of low-cardinality classification rules that has an adequate accuracy level and is easy to interpret. To achieve this, two important aspects were considered – the first of these aspects in relation to the method's ability to operate with fuzzy attributes, and the second aspect in relation to the insertion of information based on degrees of membership, both for fitness function evaluation as well as for the search process using the optimization technique. It uses a variable size particle swarm, initialized through a competitive neural network and using linguistic variables to express the conditions related to numeric variables.

The i^{th} particle in the population for the version that uses fuzzy attributes in FRvarPSO is represented as follows:

- $pBin_i = (pBin_{i1}, pBin_{i2}, \ldots, pBin_{in})$ is a binary vector that stores the current position of the particle and indicates which are the items or conditions that form the antecedent of the rule according to PSO.
- $v1_i = (v1_{i1}, v1_{i2}, \ldots, v1_{in})$ and $v2_i = (v2_{i1}, v2_{i2}, \ldots, v2_{in})$ are combined to determine the direction in which the particle will move.
- $pBestBin_i = (pBestBin_{i1}, pBestBin_{i2}, \ldots, pBestBin_{in})$ stores the best solution found for the particle so far.
- $fitness_i$ is the fitness value for the individual.
- $fitness\_pBest_i$ is the fitness value for the best local solution found ($pBestBin_i$ vector)
- $L_i = (L_{i1}, L_{i2}, \ldots, L_{in})$ is a binary vector that indicates the possible values that each fuzzy variable can have.
- $Gp_i = (Gp_{i1}, Gp_{i2}, \ldots, Gp_{in})$ is a real values vector that stores the average degrees of membership corresponding to the examples that fulfill the rule for each value of the fuzzy attributes, each of which will be represented by three fuzzy sets obtained through the Fuzzy C-means algorithm.
- $v3_i = (v3_{i1}, v3_{i2}, \ldots, v3_{in})$ indicates the change direction for L_i, with degree of membership Gp_i.
- $pBestGp_i = (pBestGp_{i1}, pBestGp_{i2}, \ldots, pBestGp_{in})$ stores the best solution found by the particle for the degrees of membership corresponding to the linguistic variable.
- $sopBin_i = (sopBin_{i1}, sopBin_{i2}, \ldots, sopBin_{in})$ indicates which are the items or conditions that form the antecedent of the rule that effectively represents the particle and whose fitness is in $fitness_i$.
- *TV* is an integer that indicates the remaining life time for the particle. It is used only when working with variable population sizes.

The movement of the i^{th} particle is controlled through a variation of PSO directed by velocity vectors $v1_i$ and $v2_i$, with $pBin_i$ being the result of applying the sigmoid function. The binary individual that selects the conditions included in the antecedent of the rule is expressed in $sopBin_i$, and it comes from $pBin_i$ after removing all invalid solutions. To decide the value with which each fuzzy variable will be able to participate in the condition, vector Gp_i is added, with the average degrees of membership of each fuzzy variable in the different sets. This average is calculated by considering the degrees of membership for the examples that fulfill the antecedent of the rule when fitness is evaluated. This vector is the one used to modify velocity vector $v3_i$. L_i is the result of applying the sigmoid function to $v3_i$ and, therefore, is a binary vector that indicates the possible values that each fuzzy variable can have if it is selected.

The fitness function is defined as follows:

$$\text{Fitness} = (\text{support} * \text{confidence} * \text{factor1}) - (\text{factor2} * \text{NumAttribs Antecedent}) / \text{MaxAttribs} \tag{5}$$

Factor1 is a penalization value for cases where the support is outside the ranges established in the algorithm. The second term in the fitness function reflects the importance given to the number of attributes included in the antecedent, factor2 being a constant.

Once the first rule has been obtained, the degree of membership for each of the examples that fulfill the conditions of rule Gp_{i}, is also obtained, where Gp_{i}, is the degree of membership of example *i* that fulfills the conditions of the rule, which is given by a t-norm that uses the minimum operator between the degrees of membership of the fuzzy attributes involved in the antecedent of the rule.

Then, voting criterion CV is calculated, which is the average of all degrees of memberships for the examples that fulfill the rule, i.e., is given by:

$$CV = \frac{\sum_{i=1}^{n} G_{pi}}{n} \tag{6}$$

This voting criterio CV is used for fuzzy attribute selection, and is considered for calculating the movement of individuals by adding it directly to the corresponding velocity vector. This favors selecting the attributes with higher voting criterion values. This method is described in more detail in [2].

Algorithm 1 shows the pseudocode used.

Algorithm 1: Pseudocode for Fuzzy C-Means and FRvarPSO

For each numeric variable
  Choose the number of clusters C
  Choose the value of the m
  Choose the similarity measure (use Euclidean distance)
  Assign tolerance $\varepsilon \leq 0.01$
  Generate the matrix for the degrees of membership μ_{ij} n a random manner
    While *(The change in the position of the centers is greater than ε)*
      Calculate new centers using equation (3)
      Update the matrix for the degrees of membership μ_{ij} following equation (2)
    End while
Train the competitive neural network using the training examples
Determine the minimum support for each class
While *(there are enough uncovered examples)*
  Choose the class with the highest number of examples that have not been covered
  Build a population considering neural network centroids
  Evaluate the fitness value of each particle using equation (5)
  While *the particle population does not reach a stable status*
    Identify the best solution found so far
    For each *particle*
      Calculate the voting criterion (average degree of membership of the examples that meet the rule indicated by the particle) using equation (6)
      Calculate the speed and add it to the vote criterion mentioned above
      Obtain the new position for the particle by adding the speed mentioned above and limit as appropriate
    End for
    If using elitism, recover the best solution
  End while
  Obtain the best fuzzy rule for the population
  If *the fuzzy rule meets support and confidence requirements,* ***then***
    Add the fuzzy rule to the fuzzy rule set
    Obtain the output value for the fuzzy rule, given by the degree of membership of the examples that fulfill the rule, using the corresponding t-norm.
    Remove from the input set those examples that are correctly covered
    Recalculate the minimum support for the class that has been considered
  End If
End while

IV. Data and Results Obtained

In this section, a performance comparison is described for the method proposed combining Fuzzy C-means and FRvarPSO, obtaining fuzzy classification rules that have fuzzy and/or nominal variables, based on combining clustering, competitive neural networks (LVQ) and variable population optimization techniques (varPSO).

To verify the performance of this method, twelve databases from the UCI repository [27], one database from a credit and savings cooperative from the Ecuadorian financial system and two databases from two Ecuadorian banks – one that grants productive and non-productive consumption credits and one that grants microcredits – were used. Thirty separate instances were run for each method; the LVQ used consists of 30 neurons.

Table I shows the results obtained for accuracy when applying four methods that are described [2], [25], [26], for the first case, an equitable distribution of the fuzzy sets was carried out, and in the second case, Fuzzy C-means was used for obtaining the fuzzy sets; the corresponding standard deviations are also listed.

Table II shows the results obtained for Number of rules when applying four methods that are described [2], [25], [26], for the first case, an equitable distribution of the fuzzy sets was carried out, and in the second case, Fuzzy C-means was used for obtaining the fuzzy sets; the corresponding standard deviations are also listed.

Table III shows the results compared accuracy and number of rules results for the database from the credit and savings cooperative in Ecuador, the Table IV shows comparative results for Accuracy and Number of Rules for a database from a bank in Ecuador that grants microcredits and is specialized in mass credit placement, and the Table V shows Comparative results for Accuracy and Number of Rules for a bank that grants consumption and productive or business credits.

TABLE I. ACCURACY RESULTS, LISTING THE BEST SOLUTION OBTAINED WITH OTHER METHODS COMPARED TO FRVARPSO WITH EQUITABLE FUZZY SETS VS. FUZZY SETS WITH FUZZY C-MEANS, ALSO SHOWING THE CORRESPONDING STANDARD DEVIATION VALUES

Data Set	**Best solution with other methods [2], [25], [26]**		**FRvarPSO {25]**	**FRvarPSO FCM**
Adult_data	SOM+Fuzzy PSO FCM	**0.8501 (0.0020)**	0.8118 (0.0047)	**0.8326 (0.0066)**
Balance_scale	SOM + Fuzzy varPSO FCM	0.7485 (0.0152)	0.7619 (0.0015)	**0.7761 (0.0020)**
Breast_w	Fuzzy LVQ + PSO	0.9544 (0.0114)	0.9565 (0.0063)	**0.9682 (0.0033)**
Credit_a	SOM + Fuzzy varPSO FCM	**0.8838 (0.0052)**	0.8689 (0.0122)	**0.8738 (0.0041)**
Credito_coop	LVQ + Fuzzy PSO FCM	0.7890 (0.0035)	0.7815 (0.0041)	**0.7901 (0.0092)**
Credit_g	Fuzzy SOM + varPSO	**0.7697 (0.0081)**	0.7592 (0.0058)	**0.7677 (0.0326)**
Credit_banco1	LVQ + Fuzzy PSO FCM	0.9811 (0.0032)	0.9551 (0.0057)	**0.9872 (0.0005)**
Credit_banco2	LVQ + Fuzzy PSO FCM	0.8410 (0.0033)	0.8423 (0.0027)	**0.8488 (0.0092)**
Diabetes	LVQ + Fuzzy PSO FCM	0.7402 (0.0182)	0.7442 (0.0148)	**0.7494 (0.0169)**
Drugte	LVQ + Fuzzy PSO FCM	0.8575 (0.0172)	0.8513 (0.0195)	**0.8595 (0.0204)**
Heart_c	Fuzzy LVQ + PSO	**0.7983 (0.0024)**	0.7866 (0.0248)	**0.7933 (0.0329)**
Heart_Statlog	SOM + Fuzzy varPSO FCM	0.7925 (0.0125)	0.7962 (0.0236)	**0.8014 (0.0254)**
Vinos	LVQ + Fuzzy PSO FCM	**0.9005 (0.0065)**	0.8944 (0.0109)	**0.8999 (0.0406)**
Zoo	LVQ + Fuzzy PSO FCM	0.9600 (0.0099)	0.9601 (0.0023)	**0.9699 (0.0045)**

TABLE II. ACCURACY RESULTS, LISTING THE BEST SOLUTION OBTAINED WITH OTHER METHODS COMPARED TO FRVARPSO WITH EQUITABLE FUZZY SETS VS. FUZZY SETS WITH FUZZY C-MEANS, ALSO SHOWING THE CORRESPONDING STANDARD DEVIATION VALUES

Data Set	**Best solution with other methods [2], [25], [26]**		**FRvarPSO {25]**	**FRvarPSO FCM**
Adult_data	Fuzzy SOM + varPSO	**2.3466 (0.2828)**	3.0106 (0.2267)	**2.2366 (0.2580)**
Balance_scale	SOM + Fuzzy varPSO FCM	7.3300 (0.2141)	8.1523 (0.5794)	**7.1200 (0.2017)**
Breast_w	SOM + Fuzzy varPSO FCM	2.3500 (0.1670)	**2.3256 (0.1567)**	2.4311 (0.1400)
Credit_a	SOM + Fuzzy varPSO FCM	2.9800 (0.0010)	3.0000 (0.0005)	**2.8055 (0.0022)**
Credito_coop	LVQ + Fuzzy PSO FCM	5.7600 (0.1453)	5.7321 (0.1357)	**5.4081 (0.1275)**
Credit_g	SOM + Fuzzy PSO FCM	7.5401 (0.0115)	8.2120 (0.0564)	7.9602 (0.0206)
Credit_banco1	SOM + Fuzzy varPSO FCM	6.6220 (0.2811)	6.9086 (0.2011)	**6.5901 (0.1723)**
Credit_banco2	SOM + Fuzzy PSO FCM	6.4911 (0.1712)	6.8716 (0.3018)	**6.3102 (0.2872)**
Diabetes	SOM + Fuzzy varPSO FCM	4.1732 (0.1484)	**4.1371 (0.1647)**	**4.1020 (0.1523)**
Drugte	LVQ + Fuzzy PSO FCM	7.5300 (0.1525)	7.2002 (0.3466)	**7.0501 (0.1250)**
Heart_c	LVQ + Fuzzy PSO FCM	3.400 (0.0221)	3.5908 (0.0129)	**3.2300 (0.0132)**
Heart_Statlog	LVQ + Fuzzy PSO FCM	3.4289 (0.0121)	3.3468 (0.0316)	**3.1295 (0.0124)**
Vinos	LVQ + Fuzzy PSO FCM	4.2200 (0.0105)	4.1144 (0.0481)	**4.0009 (0.0327)**
Zoo	LVQ + Fuzzy PSO FCM	6.7002 (0.0098)	6.5422 (0.0156)	**6.3107 (0.0107)**

TABLE III. COMPARED ACCURACY AND NUMBER OF RULES RESULTS FOR THE DATABASE FROM THE CREDIT AND SAVINGS COOPERATIVE IN ECUADOR, USING EQUITABLE DISTRIBUTION FOR FUZZY SETS, FUZZY-C MEANS, AND EXPERT CRITERION

Method	Prediction Type	Rejected	Granted	Accuracy	# of rules
FRvarPSO Equitable Sets	Rejected	**0.6120** ±0.0034	0.0865 ±0.0030	0.7815 ±0.0041	5,7321 ±0.1357
	Granted	0.1319 ±0.0027	**0.1695** ±0.0022		
FRvarPSO FCM	Rejected	**0.6196** ±0.0035	0.0857 ±0.0022	0.7901 ±0.0092	5.4081 ±0.1275
	Granted	0.1241 ±0.0039	**0.1705** ±0.0027		
FRvarPSO Expert	Rejected	**0.6499** ±0.0021	0.1145 ±0.0017	**0.7988** ±0.0029	**5.2990** ±0.1907
	Granted	0.0867 ±0.0016	**0.1489** ±0.0028		

TABLE IV. COMPARATIVE RESULTS FOR ACCURACY AND NUMBER OF RULES FOR A DATABASE FROM A BANK IN ECUADOR THAT GRANTS MICROCREDITS, USING EQUITABLE DISTRIBUTION FOR FUZZY SETS, FUZZY C-MEANS, AND EXPERT CRITERION

Method	Prediction Type	Rejected	Granted	Accuracy	# of rules
FRvarPSO Equitable Sets	Rejected	**0.8623** ±0.0029	0.0376 ±0.0017	0.9551 ±0.0057	6.9086 ±0.2011
	Granted	0.0072 ±0.0019	**0.0928** ±0.0022		
FRvarPSO FCM	Rejected	**0.8831** ±0.0038	0.0071 ±0.0041	0.9872 ±0.0005	6.5901 ±0.1723
	Granted	0.0056 ±0.0029	**0.1041** ±0.0045		
FRvarPSO Expert	Rejected	**0.9037** ±0.0023	0.0089 ±0.0024	**0.9880** ±0.0026	**6.3972** ±0.1915
	Granted	0.0031 ±0.0027	**0.0843** ±0.0029		

TABLE V. COMPARATIVE RESULTS FOR ACCURACY AND NUMBER OF RULES FOR A BANK THAT GRANTS CONSUMPTION AND PRODUCTIVE OR BUSINESS CREDITS, USING EQUITABLE DISTRIBUTION FOR FUZZY SETS, FUZZY C-MEANS, AND EXPERT CRITERION

Method	Prediction Type	Rejected	Granted	Accuracy	# of rules
FRvarPSO Equitable Sets	Rejected	**0.6728** ±0.0022	0.0981 ±0.0017	0.8423 ±0.0027	6.8716 ±0.3018
	Granted	0.0595 ±0.0019	**0.1695** ±0.0020		
FRvarPSO FCM	Rejected	**0.6818** ±0.0029	0.0807 ±0.0041	0.8488 ±0.0092	6.3102 ±0.2872
	Granted	0.0704 ±0.0032	**0.1670** ±0.0037		
FRvarPSO Expert	Rejected	**0.6627** ±0.0034	0.0830 ±0.0028	**0.8501** ±0.0033	**5.9901** ±0.3112
	Granted	0.0669 ±0.0030	**0.1874** ±0.0022		

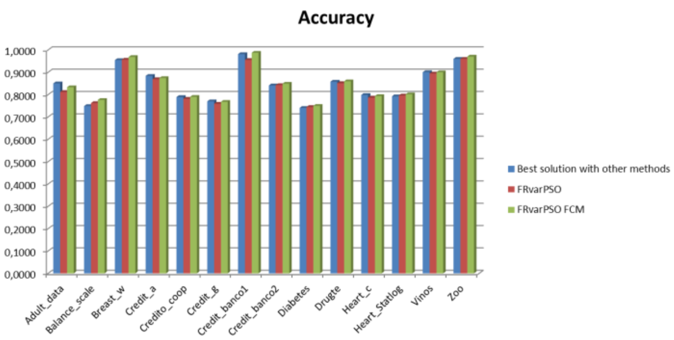

Figure 1. Accuracy results obtained with the best accuracy for each of the databases analyzed.

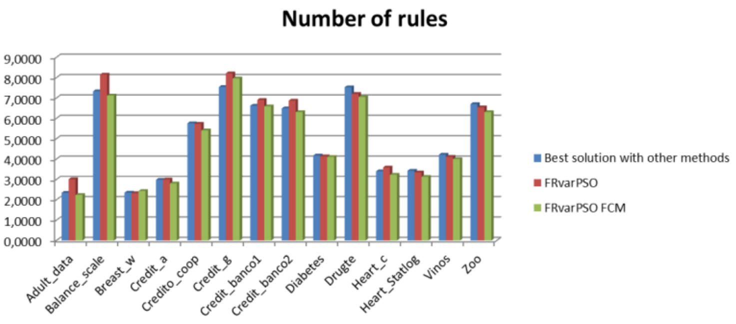

Figure 2. Number of rules results obtained with the lowest number of rules for each of the databases analyzed.

Fig. 1 shows Accuracy results obtained with the best accuracy for each of the databases analyzed, and the Fig. 2 number of Rules results obtained with the lowest number of rules for each of the databases analyzed, where you can see how FRvarPSO FCM gives excellent results.

V. Conclusions and Future Work

We have presented the FRvarPSO method, which uses a neural network with LVQ supervised learning and a variable population metaheuristic technique varPSO, applying a voting criterion for particle movement, obtaining fuzzy and/or nominal attributes in the antecedent of the rule used to generate the fuzzy rules. In the first case, to establish the degree of membership of each variable to the fuzzy sets, equitable distribution was used to identify fuzzy sets when obtaining the fuzzy variables, considering uniform fuzzy partitions with the same number of labels for each variable. In the second case, the Fuzzy C-means algorithm was used with 3 clusters, each of them representing a fuzzy set.

Measurements show that using hybrid algorithms such as Fuzzy C-means and FRvarPSO allows improving accuracy and obtaining a simpler model with less rules, while achieving a better interpretation ability by the user, with no need to resort to an expert.

In the future, a larger number of clusters should be considered for establishing the degree of membership for input variables, as well as considering variations in the FCM algorithm to work with inaccurate data. Other variations can also be considered, such as establishing the consequent of the rule as a fuzzy variable, using the voting criterion, which can provide a degree of certainty in the result obtained, with the corresponding decrease in risk when making a decision.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Laura Lanzarini headed the research line that supported this article, Patricia Jimbo made the corresponding measurements and wrote the article. Aurelio Fernández designed the experiments carried out and supervised the results obtained. All authors had approved the final version.

References

[1] C. Aggarwal, *Data Mining: The Textbook*, Springer Publishing Company, 2015.

[2] P. J. Santana, L. Lanzarini, and A. F. Bariviera, "FRvarPSO: A method for obtaining fuzzy classification rules using optimization techniques," *Modelling and Simulation in Management Sciences*, vol. 894, pp. 112-126, 2019.

[3] L. Wang and J. M. Mendel, "Generating fuzzy rules by learning from examples," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 22, no. 6, pp. 1414-1427, Nov.-Dec. 1992.

[4] S. Yong and L. Bongju, "The one-sided quadrangular fuzzy sets," *Journal of the Chungcheong Mathematical Society*, vol. 26, no. 2, pp. 297-308, May 2013.

[5] J. Barrios, C. Villanueva, A. Cavazos, and R. Colas, "Fuzzy C-means rule generation for fuzzy entry temperature prediction in a hot strip mill," *Journal of Iron and Steel Research, International*, vol. 23, no. 2, pp. 116-123, 2016.

[6] R. Yager and P. Filev, "Generation of fuzzy rules by mountain clustering," *Journal of Intelligent and Fuzzy Systems*, vol. 2, no. 3, pp. 209-219, 1994.

[7] S. L. Szilágyi and M. Szilágyi, "Generalization rules for the suppressed fuzzy c-means clustering algorithm," *Neurocomputing*, vol. 139, pp. 298-309, 2014.

[8] H. Roubos and M. Setnes, "Compact fuzzy models through complexity reduction and evolutionary optimization," in *Proc. IEEE International Conference on Fuzzy Systems*, 2000.

[9] H. Ishibuchi and T. Yamamoto, "Rule weight specification in fuzzy rule-based classification systems," *IEEE Transactions on Fuzzy Systems*, vol. 13, no. 4, pp. 428-435, 2005.

[10] Y. Chung, R. Shun, and G. Hshiung, "Finding fuzzy classification rules using data mining techniques," *Pattern Recognition Letters*, vol. 24, no. 1, pp. 509-519, 2003.

[11] G.. Soumadip, B. Sushanta, D. Sarkar, and P. Pratim, "A novel neuro-fuzzy classification technique for data mining," *Egyptian Informatics Journal*, vol. 15, pp. 129-147, 2014.

[12] X. Hu, W. Pedrycz, and X. Wang, "Random ensemble of fuzzy rule-based models," *Knowledge-Based Systems*, vol. 181, p. 104768, 2019.

[13] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 1, pp. 116-132, 1985.

[14] W. Pedrycz and M. Reformat, "Evolutionary fuzzy modeling," *IEEE Transactions on Fuzzy Systems*, vol. 11, pp. 652-665, 2003.

[15] M. Ahmed and N. Isa, "Knowledge base to fuzzy information granule: A review from the interpretability-accuracy perspective," *Applied Soft Computing*, vol. 54, pp. 121-140, May 2017.

[16] X. Hu, W. Pedrycz, and X. Wang, "Granular fuzzy rule-based models: A study in a comprehensive evaluation of fuzzy models," *IEEE Transactions on Fuzzy Systems*, vol. 25, pp. 1342-1355, 2017.

[17] X. Hu, W. Pedrycz, and X. Wang, "Development of granular models through the design of a granular output spaces," *Knowledge-Based Systems*, vol. 134, pp. 159-171, 2017.

[18] L. Zhang, M. Luo, J. Liu, Z. Li, and Z. Qingh, "Diverse fuzzy c-means for image clustering," *Pattern Recognition Letters*, 2018.

[19] R. Jin and G. Weng, "A robust active contour model driven by fuzzy c-means energy for fast image segmentation," *Digital Signal Processing*, vol. 90, pp. 100-109, 2019.
[20] J. Xu, G. Feng, T. Zhao, X. Sun, and M. Zhu, "Remote sensing image classification based on semi-supervised adaptive interval type-2 fuzzy C-means algorithm," *Computers & Geosciences*, vol. 131, pp. 132-143, 2019.
[21] A. Stetco, X. Jun, and Z. J. Keane, "Fuzzy C-means++: Fuzzy C-means with effective seeding initialization," *Expert Systems with Applications*, vol. 42, pp. 7541-7548, 2015.
[22] B. Kindhi, T. Arief, M. Hery, and J. Gijbertus, "Hybrid K-means, fuzzy C-means, and hierarchical clustering for DNA hepatitis C virus trend mutation analysis," *Expert Systems with Applications*, vol. 121, pp. 373-381, 2019.
[23] E. Hanyu, C. Ye, P. Witold, and Z. Li, "Enhancements of rule-based models through refinements of fuzzy C-means," *Knowledge-Based Systems*, vol. 170, pp. 43-60, 2019.
[24] C. Bezdek, R. Ehrlich, and W. Full, "FCM: The fuzzy C-means clustering algorithm," *Computers & Geosciences*, vol. 10, pp. 191-203, 1984.
[25] P. J. Santana, L. Lanzarini, and A. F. Bariviera, "Fuzzy credit risk scoring rules using FRvarPSO," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 26, suppl. 1, pp. 39-57, 2018.
[26] P. J. Santana, L. Lanzarini, and A. F. Bariviera, "Extraction of knowledge with population-based metaheuristics fuzzy rules applied to credit risk," *Advances in Swarm Intelligence*, pp. 153-163, 2018.
[27] UCI Irvine machine learning repository. [Online]. Available: http://archive.ics.uci.edu/ml/

Patricia Jimbo Santana is Full Professor at the Central University of Ecuador, is an Engineer in Computer and Computer Systems, Computer Expert of the Criminology Institute of the Central University of Ecuador. She is doing her PhD in Computer Science at the National University of La Plata, Argentina, in Cotutela with the University of Virgina Rovaire of Spain, among her research lines are data mining, machine learning, risk, information and communication technologies.

Her research interests include data mining, machine learning, information and communication technologies and risks.

Laura Lanzarini is Full Professor at La Plata University in Buenos Aires, Argentina.

She is Doctor of Computer Science and leads a research group dedicated to the development of Intelligent Systems.

Her research interests include data mining, machine learning, big data and pattern recognition.

Aurelio F, Bariviera is Associate Professor at Universitat Rovira i Virgili, Spain.

He holds a PhD in Economics and his research is devoted to complex systems, data mining, forecasting and econophysics. He has published previously in Physica A, Economics Letters, and Philosophycial Transactions of the Royal Societe, among other journals.