Improving the performance of fuzzy rule-based classification systems based on a non-averaging generalization of CC-integrals named $C_{F_1F_2}$-integrals

Giancarlo Lucca, Graçaliz Pereira Dimuro *IEEE Member*, Javier Fernández, Humberto Bustince, *IEEE Senior Member*, Benjamín Bedregal *IEEE Member*, José Antonio Sanz

Abstract—A key component of Fuzzy Rule-Based Classification Systems (FRBCSs) is the Fuzzy Reasoning Method (FRM), since it infers the class predicted for new examples. A crucial stage in any FRM is the way in which the information given by the fired rules during the inference process is aggregated. A widely used FRM is the winning rule, which applies the maximum to accomplish this aggregation. The maximum is an averaging operator, which means that its result is within the range delimited by the minimum and the maximum of the aggregated values. Over the last years, new averaging operators based on generalizations of the Choquet integral were also proposed to perform this aggregation process. However, the most accurate FRBCSs use the FRM known as additive combination, that considers the normalized sum as the aggregation operator, which is non-averaging. For this reason, this paper is aimed at introducing a new non averaging operator named $C_{F_1F_2}$-integral, which is a generalization of the Choquet-like Copula-based integral (CC-integral). $C_{F_1F_2}$-integrals present the desired properties of an aggregation-like operator, since they satisfy appropriate boundary conditions and have some kind of increasingness property. We show that $C_{F_1F_2}$-integrals, when used to cope with classification problems, enhance the results of the previous averaging generalizations of the Choquet integral and they provide competitive results (even better) when compared with state-of-the-art FRBCSs.

Index Terms—Fuzzy rule-based classification systems, Choquet Integral, $C_{F_1F_2}$-integrals, CC-integrals, OD monotone functions.

I. INTRODUCTION

In a supervised classification problem [1] it is necessary to determine the class of an example based on the information given by labeled examples. Among others, an accurate way to tackle classification problems is by using Fuzzy Rule-Based Classification Systems (FRBCSs) [2]. This technique achieves accurate results taking into consideration linguistic labels in

G. Lucca is with Departamento of Automática y Computación, Universidad Publica de Navarra, Navarra, 31006 Spain e-mail: lucca.112793@e.unavarra.es

H. Bustince, J. Fernández and J. Sanz are with the Departamento de Automática y Computación and with the Institute of Smart Cities, Universidad Publica de Navarra, Navarra, 31006 Spain e-mails: {bustince,fcojavier.fernandez,joseantonio.sanz,}@unavarra.es.

G. Pereira Dimuro is with Institute of Smart Cities, Universidad Publica de Navarra, Pamplona, Spain, and Centro de Ciências Computacionais, Universidade Federal do Rio Grande, Rio Grande, Brazil, e-mail: gracaliz dimuro@furg.br.

B. Bedregal is with Departamento de Informática e Matemática Aplicada, Universidade Federal do Rio Grande do Norte, Natal, Brazil, e-mail: bedregal@dimap.ufrn.br.

the rules, which leads to obtaining an interpretable model that can be easily used in the decision making process.

The two main components of FRBCSs are the knowledge base, which is composed of the rule base and the data base, and the Fuzzy Reasoning Method (FRM) [3]. The latter is a mechanism that uses the information available in the knowledge base to assign a class to new examples that have to be classified. The FRM of the winning rule is a classical inference process found in the literature that assigns the class of the fuzzy rule whose compatibility with the example to be classified is maximum. To do so, it applies the maximum as the aggregation operator, which has an averaging characteristic, i.e., its result is delimited by the minimum and the maximum of the values to be aggregated. Consequently, to classify an example it only uses the information given by one fuzzy rule and it disregards the remainder information.

Barrenechea et. al proposed in [4] a FRM that takes into account the information provided by all the fired rules using the Choquet integral [5]. After that, the Choquet integral was generalized by replacing the standard product operator by different t-norms, which led to the concept of pre-aggregation functions [6]. Next, aiming at producing an aggregation function, in [7] the authors presented the Choquet-like Copula-based integral (CC-integrals, for short). They swapped the product operator of the extended form of the Choquet integral by two identical copulas C. These three approaches have averaging characteristics [8] and they provide competitive results in classification problems.

However, the state of the art FRBCSs algorithms, like IVTURS [9], FARC-HD [10] or FURIA [11] apply the FRM known as additive combination that is based on the usage of the normalized sum as aggregation function [3], which has a non-averaging behavior. Taking this fact into account, in [12] the standard Choquet integral was generalized by replacing the product operation by different functions F, introducing the concept of C_F-integrals. These integrals are pre-aggregations that may have either averaging or non-averaging characteristics according to the considered function F. The authors showed that the non-averaging C_F-integrals statistically overcome the averaging ones, reinforcing the quality of the usage of non-averaging functions in this domain.

For this reason, in this paper we define a generalization of the CC-integral, named $C_{F_1F_2}$-integral, substituting the copula C by two fusion functions F_1 and F_2 satisfying some special conditions. The $C_{F_1F_2}$-integrals are non-averaging

Ordered Directionally (OD) increasing functions satisfying the required boundary conditions for any "aggregation-like operator". Moreover, we present a methodology to select the best pairs of fusion functions F_1 and F_2 to define the $C_{F_1F_2}$-integrals to tackle classification problems. Finally, we also introduce a new FRM based on this concept.

In the experimental study, we consider 33 different datasets available in Keel dataset repository [13]. We analyze the quality of our method by selecting the best $C_{F_1F_2}$-integrals and comparing them against the state-of-the-art fuzzy classifiers, the best C_F-integral presented in [12] and the classical probabilistic sum [14] applied in the FRM of a FRBCS as it is a known non-averaging function. The quality of the results is supported by a proper statistical study as suggested in the specialized literature [15], [16], [17].

The paper is organized as follows. In Section II, we introduce the background necessary to understand the paper. In Section III, we introduce the concept of $C_{F_1F_2}$-integrals, showing that these functions are OD increasing functions satisfying appropriate conditions. We describe in Section IV the use of $C_{F_1F_2}$-integrals in the FRM and the evolutionary learning of the fuzzy measure. Section V introduces the experimental framework, describing the datasets along with the setup configuration for the different methodologies and the statistical tests used for performance comparison. In Section VI we present the experimental results achieved in testing by $C_{F_1F_2}$-integrals and we draw the main conclusions in Section VII.

II. PRELIMINARIES

In this section, we present some basic theoretical concepts that are necessary to develop the paper. Any n-ary function $F : [0,1]^n \to [0,1]$ is named fusion function as it receives n values and it fuses them returning a single one.

Definition 1. *[18], [19] A fusion function $A : [0,1]^n \to [0,1]$ is an aggregation function whenever the following conditions hold:*

(A1) *A is increasing[1] in each argument: for each $i \in \{1, \ldots, n\}$, if $x_i \leq y$, then $A(x_1, \ldots, x_n) \leq A(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n)$;*

(A2) *A satisfies the boundary conditions: (i) $A(0, \ldots, 0) = 0$ and (ii) $A(1, \ldots, 1) = 1$.*

An aggregation function $A : [0,1]^n \to [0,1]$ is said to be averaging if and only if: **(AV)** $\forall (x_1, \ldots, x_n) \in [0,1]^n : \min\{x_1, \ldots, x_n\} \leq A(x_1, \ldots, x_n) \leq \max\{x_1, \ldots, x_n\}$.

Definition 2. *[20] Let $\vec{r} = (r_1, \ldots, r_n)$ be a real n-dimensional vector, $\vec{r} \neq \vec{0}$. A function $F : [0,1]^n \to [0,1]$ is said to be $\vec{r}$-increasing if for all $\vec{x} = (x_1, \ldots, x_n) \in [0,1]^n$ and for all $c > 0$ such that $\vec{x} + c\vec{r} = (x_1 + cr_1, \ldots, x_n + cr_n) \in [0,1]^n$ it holds*

$$F(\vec{x} + c\vec{r}) \geq F(x_1, \ldots, x_n). \qquad (1)$$

Similarly, one defines an $\vec{r}$-decreasing function.

[1]For an increasing (decreasing) function we do not mean a strictly increasing (decreasing) function.

Definition 3. *[6] A function $PA : [0,1]^n \to [0,1]$ is said to be an n-ary pre-aggregation function if the following conditions hold:*

(PA1) *Directional increasingness: there exists $\vec{r} = (r_1, \ldots, r_n) \in [0,1]^n$, $\vec{r} \neq \vec{0}$, such that PA is $\vec{r}$-increasing;*

(PA2) *Boundary conditions: (i) $PA(0, \ldots, 0) = 0$ and (ii) $PA(1, \ldots, 1) = 1$.*

If F is a pre-aggregation function with respect to a vector $\vec{r}$ we just say that F is an $\vec{r}$-pre-aggregation function.

In what follows, denote $N = \{1, \ldots, n\}$ for an arbitrary $n > 0$.

Definition 4. *[5], [21][22, Definition 1.77] A function $\mathfrak{m} : 2^N \to [0,1]$ is said to be a fuzzy measure if, for all $X, Y \subseteq N$, the following conditions hold:*

($\mathfrak{m}$1) *Increasingness: if $X \subseteq Y$, then $\mathfrak{m}(X) \leq \mathfrak{m}(Y)$;*

($\mathfrak{m}$2) *Boundary conditions: $\mathfrak{m}(\emptyset) = 0$ and $\mathfrak{m}(N) = 1$.*

A fuzzy measure $\mathfrak{m}$ is symmetric whenever $\mathfrak{m}(X) = \mathfrak{m}(Y)$ for all $X, Y \subseteq N$ such that $| X | = | Y |$.

Definition 5. *[5] Let $\mathfrak{m} : 2^N \to [0,1]$ be a fuzzy measure. The discrete Choquet integral is the function $\mathfrak{C}_{\mathfrak{m}} : [0,1]^n \to [0,1]$, defined, for all of $\vec{x} = (x_1, \ldots, x_n) \in [0,1]^n$, by:*

$$\mathfrak{C}_{\mathfrak{m}}(\vec{x}) = \sum_{i=1}^{n} \left(x_{(i)} - x_{(i-1)} \right) \cdot \mathfrak{m} \left(A_{(i)} \right), \qquad (2)$$

where $\left(x_{(1)}, \ldots, x_{(n)} \right)$ is an increasing permutation on the input $\vec{x}$, that is, $0 \leq x_{(1)} \leq \ldots \leq x_{(n)}$, where $x_{(0)} = 0$ and $A_{(i)} = \{(i), \ldots, (n)\}$ is the subset of indices corresponding to the $n - i + 1$ largest components of $\vec{x}$.

The CC-integral [7] is a generalization of the Choquet integral using copulas [14].

Definition 6. *[7, Definition 7] Let $\mathfrak{m} : 2^N \to [0,1]$ be a fuzzy measure and $C : [0,1]^2 \to [0,1]$ a copula. The CC-integral with respect to $\mathfrak{m}$ is the function $\mathfrak{C}_{\mathfrak{m}}^C : [0,1]^n \to [0,1]$, defined, for all of $\vec{x} = (x_1, \ldots, x_n) \in [0,1]^n$, by:*

$$\mathfrak{C}_{\mathfrak{m}}^C(\vec{x}) = \sum_{i=1}^{n} \left(C \left(x_{(i)}, \mathfrak{m} \left(A_{(i)} \right) \right) - C \left(x_{(i-1)}, \mathfrak{m} \left(A_{(i)} \right) \right) \right). \qquad (3)$$

III. A GENERALIZATION OF CC-INTEGRALS USING TWO FUSION FUNCTIONS F_1 AND F_2

In this section, we introduce a method for constructing a generalization of CC-integrals, named $C_{F_1F_2}$-integral, using two fusion functions F_1 and F_2 satisfying some specific properties instead of the same copula C (Section III.A). We also present a mechanism for choosing the functions F_1 and F_2 that will be applied in the FRM of FRBCSs (Section III.B). Finally, we prove that $C_{F_1F_2}$-integrals built with some specific pairs of fusion functions F_1 and F_2 are non-averaging OD increasing functions satisfying proper boundary conditions to be applied in the FRM (Section III.C).

A. Defining the $C_{F_1F_2}$-integrals

In this subsection, we aim at introducing the definition of $C_{F_1F_2}$-integrals and analyzing some properties for specific pairs of fusion functions F_1 and F_2.

An important concept used in this paper is the dominance (or, conversely, subordination) property:

(DM) F_1-Dominance (or, equivalently, F_2-Subordination): $F_1 \geq F_2$, that is: $\forall$ x,y $\in [0,1]$: F_1 (x,y) $\geq F_2$(x,y)

Definition 7. *Let* $\mathfrak{m} : 2^N \rightarrow [0,1]$ *be a symmetric fuzzy measure and* $F_1, F_2 : [0,1]^2 \rightarrow [0,1]$ *be two fusion functions fulfilling:*

(i) F_1*-dominance*
(ii) F_1 *is* $(1,0)$*-increasing,*

A $C_{F_1F_2}$*-integral is defined as a function* $\mathfrak{C}_{\mathfrak{m}}^{(F_1,F_2)} : [0,1]^n \rightarrow [0,1]$*, given, for all* $x \in [0,1]^n$*, by*

$$\mathfrak{C}_{\mathfrak{m}}^{(F_1,F_2)}(\vec{x}) = \tag{4}$$
$$\min\left\{1, x_{(1)} + \sum_{i=2}^{n} F_1\left(x_{(i)}, \mathfrak{m}\left(A_{(i)}\right)\right) - F_2\left(x_{(i-1)}, \mathfrak{m}\left(A_{(i)}\right)\right)\right\},$$

where $(x_{(1)}, \ldots, x_{(n)})$ *is an increasing permutation on the input* $\vec{x}$*, that is,* $0 \leq x_{(1)} \leq \ldots \leq x_{(n)}$*, with the convention that* $x_{(0)} = 0$*, and* $A_{(i)} = \{(i), \ldots, (n)\}$ *is the subset of indices of* $n - i + 1$ *largest components of* $\vec{x}$*.*

Observe that it is immediate that $\mathfrak{C}_{\mathfrak{m}}^{(F_1,F_2)}$ is well defined, for any pair $F_1, F_2 : [0,1]^2 \rightarrow [0,1]$ and a symmetric fuzzy measure $\mathfrak{m}$.

Remark 1. *Observe that the first element of the summation in the definition of* $\mathfrak{C}_{\mathfrak{m}}^{(F_1,F_2)}$ *is just* $x_{(1)}$ *instead of*

$$F_1\left(x_{(1)}, \mathfrak{m}\left(A_{(1)}\right)\right) - F_2\left(x_{(0)}, \mathfrak{m}\left(A_{(1)}\right)\right).$$

This is considered to avoid the initial discrepant behavior of non-averaging functions in the initial phase of the aggregation process. For example, consider an unitary vector $\vec{x} = 0.9 \in [0,1]$ *and* $F_1 = F_2 = AVG$*, where* $AVG(x,y) = \frac{x+y}{2}$ *is the arithmetic mean. Then, if we included the first element in the summation of the integral the result would be:*

$$\mathfrak{C}_{\mathfrak{m}}^{(F_1,F_2)}(\vec{x}) \tag{5}$$
$$= \min\left\{1, \sum_{i=1}^{n} F_1\left(x_{(i)}, \mathfrak{m}\left(A_{(i)}\right)\right) - F_2\left(x_{(i-1)}, \mathfrak{m}\left(A_{(i)}\right)\right)\right\},$$
$$= \min\left\{1, \frac{0.9+1}{2} - \frac{0+1}{2}\right\} = 0.45 \neq 0.9.$$

Observe here the large discrepancy of the result, since one expects that the aggregated value would be 0.9*. Using our definition of* $C_{F_1F_2}$*-integral (Equation (4)) this unexpected behavior is avoided and the result is 0.9.*

Table I shows the definitions of fusion functions $F : [0,1]^2 \rightarrow [0,1]$ that are $(1,0)$-increasing (condition **(ii)** of Definition 7) and, thus, candidates to be used as F_1 and/or F_2 in the definition of $C_{F_1F_2}$-integrals. The expression of the function is introduced in the first column whereas in the second and in the third columns we show the family (or families)

TABLE I: $(1,0)$-increasing fusion functions initially considered in this study

Definition	Family	Reference
$T_M(x,y) = \min\{x,y\}$	t-norm, copula	[23]
$T_P(x,y) = xy$	t-norm, copula	[23]
$T_L(x,y) = \max\{0, x+y-1\}$	t-norm, copula	[23]
$T_{HP}(x,y) = \begin{cases} 0 & \text{if } x = y = 0 \\ \frac{xy}{x+y-xy} & \text{otherwise} \end{cases}$	t-norm, copula	[23]
$T_{DP}(x,y) = \begin{cases} x & \text{if } y = 1 \\ y & \text{if } x = 1 \\ 0 & \text{otherwise} \end{cases}$	t-norm	[23]
$O_B(x,y) = \min\{x\sqrt{y}, y\sqrt{x}\}$	overlap function, copula	[24], [25]
$O_{mM}(x,y) = \min\{x,y\}\max\{x^2,y^2\}$	overlap function, copula	[26], [27], [28], [29]
$O_\alpha(x,y) = xy(1+\alpha(1-x)(1-y))$, $\alpha \in [-1,0[\cup]0,1]$	overlap function, copula	[12], [14]
$O_{Div}(x,y) = \frac{xy+\min\{x,y\}}{2}$	overlap function, copula	[14]
$GM(x,y) = \sqrt{xy}$	overlap function	[30]
$HM(x,y) = \begin{cases} 0 & \text{if } x = 0 \text{ or } y = 0 \\ \frac{2}{\frac{1}{x}+\frac{1}{y}} & \text{otherwise} \end{cases}$	overlap function	[30]
$S(x,y) = \sin\left(\frac{\pi}{2}(xy)^{\frac{1}{4}}\right)$	overlap function	[30]
$C_F(x,y) = xy + x^2 y(1-x)(1-y)$	copula	[23], [7]
$C_L(x,y) = \max\{\min\{x, \frac{y}{2}\}, x+y-1\}$	copula	[14], [7]
$AVG(x,y) = \frac{x+y}{2}$	overlap	
$F_{RS}(x,y) = \min\left\{\frac{(x+1)\sqrt{y}}{2}, y\sqrt{x}\right\}$	aggregation function	[12]
$F_{GL}(x,y) = \sqrt{\frac{x(y+1)}{2}}$	aggregation function	[12]
$F_{BPC}(x,y) = xy^2$	aggregation function	[18]
$F_\alpha(x,y) = \begin{cases} \alpha x & \text{if } x < y \\ \max\{\alpha x, y\} & \text{otherwise} \end{cases}$, $0 < \alpha < 1$	Pre-aggregation function	[12]
$F_{NA}(x,y) = \begin{cases} x & \text{if } x \leq y \\ \min\{\frac{x}{2}, y\} & \text{otherwise} \end{cases}$	Pre-aggregation function	[12]
$F_{NA2}(x,y) = \begin{cases} 0 & \text{if } x = 0 \\ \frac{x+y}{2} & \text{if } 0 < x \leq y \\ \min\{\frac{x}{2}, y\} & \text{otherwise} \end{cases}$	Pre-aggregation function	[12]
$F_{IM}(x,y) = \max\{1-y, x\}$	Non Pre-Aggregation function	[12]
$F_{IP}(x,y) = 1-y+xy$	Non Pre-Aggregation function	[12]

* When $\alpha = 0$, we have that $O_\alpha = T_P$, the product t-norm.

of the function and the source where they were published, respectively.

As we have mentioned, all these functions fulfill condition **(ii)** of Definition 7. Therefore, we need to study whether they fulfill condition **(i)**. Consequently, we conduct the study about the dominance property in the next subsection.

B. Analyzing the Dominance (Subordination) property

In this paper, $C_{F_1F_2}$-integrals are applied in the FRM of a FRBCS (see Section IV). To select the pairs of fusion functions (F_1, F_2) to be used in the construction of the $C_{F_1F_2}$-integrals, we consider the analysis of the dominance property. This property plays then a central role in the construction of the selected $C_{F_1F_2}$-integrals discussed in this paper. We analyze such property in order to determine which fusion functions, among those presented in Table I, are more suitable to be F_1 or F_2 in the construction the $C_{F_1F_2}$-integrals.

To do so, we define the concepts of dominance and subordination strength degrees. Let $\mathcal{F} = \{F_1, \ldots, F_m\}$ be a set of m fusion functions. The dominance and subordination strength degrees, DSt and SSt, of a fusion function $F_i \in \mathcal{F}$ are defined for $j \in \{1, \ldots, m\}$ as follows:

$$DSt(F_i) = \frac{1}{m}\sum_{j=1}^{m}\begin{cases} 1 & \text{if } F_i \geq F_j, \\ 0 & \text{otherwise} \end{cases} \cdot 100\%$$

$$SSt(F_i) = \frac{1}{m}\sum_{j=1}^{m}\begin{cases} 1 & \text{if } F_i < F_j, \\ 0 & \text{otherwise.} \end{cases} \cdot 100\%$$

That is, the DSt and SSt degrees of a fusion function F take into account the number of functions in which F dominates, or is subordinated to, respectively.

Table II presents the analysis of the dominance property for the functions presented in Table I. In this table a cell is

TABLE II: Analysis of the dominance property of the fusion functions introduced in Table I

	T_P	T_M	$T_Ł$	T_{DP}	T_{HP}	O_B	O_{mM}	O_α	O_{Div}	GM	HM	S	F_{RS}	C_F	C_L	F_{GL}	F_{BPC}	F_{NA}	F_α	F_{NA2}	AVG	F_{IM}	F_{IP}	$DSt\,(\%)$
T_P	✓		✓	✓			✓										✓							21.74
T_M	✓	✓	✓	✓	✓	✓	✓	✓	✓					✓	✓		✓	✓						56.52
$T_Ł$			✓	✓																				8.70
T_{DP}				✓																				4.35
T_{HP}	✓		✓	✓	✓		✓	✓						✓			✓							34.78
O_B	✓		✓	✓		✓	✓	✓									✓							30.43
O_{mM}				✓			✓										✓							13.04
O_α	✓		✓	✓			✓	✓									✓							26.09
O_{Div}	✓		✓	✓		✓	✓	✓	✓								✓							34.78
GM	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓		✓	✓						69.57
HM	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓			✓	✓		✓	✓						60.87
S	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	100
F_{RS}	✓		✓	✓		✓	✓	✓					✓				✓							34.78
C_F	✓		✓	✓			✓							✓			✓							26.09
C_L			✓	✓											✓		✓							17.39
F_{GL}	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓					78.26
F_{BPC}				✓													✓							8.70
F_{NA}				✓														✓						8.70
F_α				✓															✓					8.70
F_{NA2}				✓														✓		✓				13.04
AVG	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓		✓	✓	✓	✓	✓			82.61
F_{IM}	✓	✓	✓	✓	✓	✓	✓	✓	✓					✓	✓		✓	✓	✓			✓		65.22
F_{IP}	✓	✓	✓	✓	✓	✓	✓	✓	✓					✓	✓		✓	✓	✓			✓	✓	69.57
$SSt\,(\%)$	60.87	30.43	69.57	95.65	34.78	43.48	65.22	52.17	34.78	13.04	17.39	0.00	17.39	39.13	34.78	4.35	73.91	39.13	21.74	8.70	4.35	8.70	4.35	

marked with the ✓ symbol when the function introduced in the row dominates the one shown in the column. Furthermore, we also show in this table the DSt of the function in the row (it conforms the last column) and the SSt of the function in the column (it conforms the last row).

Since the number of possible combinations of fusion functions, marked with ✓ in Table II, for F_1 and F_2 is too high (201 different combinations), we propose a methodology to reduce the scope of this study. We consider the DSt and SSt degrees to be Low, Medium and High when they are less than 33%, between 34% and 66% and larger than 66%, respectively. Then, we have selected three functions of each category (Low, Medium, High) for both DSt and SSt to play the role of functions F_1 and/or F_2 respectively. Observe that we selected 9 different functions considering the DSt and SSt degrees. Precisely, we selected the three functions having the lowest, medium and highest degrees in each category. The selected functions according to this methodology are presented in Table III, which imply in a total number of possible combinations of 81. However, only 51 out of these 81 functions can be used as (F_1, F_2) pairs as they have an active ✓ in Table II, which means that they are pairs of functions fulfilling the (DM) property.

C. The selected $C_{F_1F_2}$-integrals as non-averaging OD monotone functions

For the aggregation process in the FRM to be well defined it is necessary an operator that has two characteristics. First, some kind of increasingness property is required in order to guarantee that the more information is provided the higher is the aggregated value (condition **(A1)** of Def. 1 and Def. 3). Second, the aggregation operator must satisfy boundary conditions related to the domain $[0, 1]$ (condition **(A2)** of Def. 1 and Def. 3).

Our selected $C_{F_1F_2}$-integrals (Table III) satisfy the boundary conditions **(A2)** of an (pre) aggregation function. However, our

TABLE III: Summary of the adopted functions according to dominance/subordination strength degrees

Strength degree	Dominance (F_1)	Subordination (F_2)
Low	T_{DP}	S
	F_{NA}	GM
	O_B	T_M
Medium	T_{HP}	T_P
	T_M	F_{NA}
	F_{IM}	T_{HP}
High	GM	$T_Ł$
	F_{GL}	F_{BPC}
	S	T_{DP}

selected $C_{F_1F_2}$-integrals are neither increasing nor directionally increasing. Nevertheless, we have noticed that they do present some kind of increasingness property. In fact, they are Ordered Directionally (OD) monotone functions [31]. Such functions are monotonic along different directions according to the ordinal size of the coordinates of each input.

In this section, we prove such properties for our best $C_{F_1F_2}$-integral, according to the results shown in Section VI ($GM-F_{BPC}$), since the proofs for the other pairs of fusion functions F_1 and F_2, considered in this paper, are analogous. We also show that they are non-averaging functions.

Definition 8. *[31] Consider a function $F : [0, 1]^n \to [0, 1]$ and let $\vec{r} = (r_1, \dots, r_n)$ be a real n-dimensional vector, $\vec{r} \neq \vec{0}$. F is said to be ordered directionally (OD) $\vec{r}$-increasing if, for each $\vec{x} \in [0, 1]^n$, any permutation $\sigma : \{1, \dots, n\} \to \{1, \dots, n\}$ with $x_{\sigma(1)} \geq \dots \geq x_{\sigma(n)}$, and $c > 0$ such that $1 \geq x_{\sigma(1)} + cr_1 \geq \dots \geq x_{\sigma(n)} + cr_n$, it holds that $F(\vec{x} + c\vec{r}_{\sigma^{-1}}) \geq F(\vec{x})$, where $\vec{r}_{\sigma^{-1}} = (r_{\sigma^{-1}(1)}, \dots, r_{\sigma^{-1}(n)})$. Similarly, one defines an ordered directionally (OD) $\vec{r}$-decreasing function.*

Theorem 1. *For any symmetric fuzzy measure $\mathfrak{m} : 2^N \to [0, 1]$ and $k > 0$, $\mathfrak{C}_\mathfrak{m}^{(GM, F_{BPC})}$, where GM and F_{BPC} are defined in Table I, is an (OD) $(k, 0, \dots, 0)$-increasing function.*

Proof. For all $\vec{x} \in [0, 1]^n$ and permutation $\sigma : \{1, \dots, n\} \to$

$\{1, \ldots, n\}$ with $x_{\sigma(1)} \geq \ldots \geq x_{\sigma(n)}$ and $c > 0$ such that $x_{\sigma(i)} + c r_i \in [0, 1]$, for $i \in \{1, \ldots, n\}$, and $1 \geq x_{\sigma(1)} + c r_1 \geq \ldots \geq x_{\sigma(n)} + c r_n$, for $\vec{r}_{\sigma^{-1}} = (r_{\sigma^{-1}(1)}, \ldots, r_{\sigma^{-1}(n)})$, one has that:

$$
\begin{aligned}
&\mathfrak{C}_{\mathfrak{m}}^{(GM, F_{BPC})}(\vec{x} + c\vec{r}_{\sigma^{-1}}) \\
&= \min \Big\{ 1, (x_{(1)} + c \cdot r_{\sigma^{-1}(1)}) \\
&\quad + \sum_{i=2}^{n-1} \Big(\sqrt{(x_{(i)} + c \cdot r_{\sigma^{-1}(i)}) \cdot \mathfrak{m}(A_{(i)})} \\
&\qquad - (x_{(i-1)} + c \cdot r_{\sigma^{-1}(i-1)}) \cdot \mathfrak{m}(A_{(i)})^2 \Big) \\
&\quad + \sqrt{(x_{(n)} + c \cdot r_{\sigma^{-1}(n)}) \cdot \mathfrak{m}(A_{(n)})} \\
&\qquad - (x_{(n-1)} + c \cdot r_{\sigma^{-1}(n-1)}) \cdot \mathfrak{m}(A_{(n)})^2 \Big\} \\
&= \min \Big\{ 1, (x_{(1)} + c \cdot 0) \\
&\quad + \sum_{i=2}^{n-1} \Big(\sqrt{(x_{(i)} + c \cdot 0) \cdot \mathfrak{m}(A_{(i)})} \\
&\qquad - (x_{(i-1)} + c \cdot 0) \cdot \mathfrak{m}(A_{(i)})^2 \Big) \\
&\quad + \sqrt{(x_{(n)} + c \cdot k) \cdot \mathfrak{m}(A_{(n)})} \\
&\qquad - (x_{(n-1)} + c \cdot 0) \cdot \mathfrak{m}(A_{(n)})^2 \Big\} \\
&\geq \min \Big\{ 1, x_{(1)} + \\
&\quad \sum_{i=2}^{n-1} \Big(\sqrt{x_{(i)} \cdot \mathfrak{m}(A_{(i)})} - x_{(i-1)} \cdot \mathfrak{m}(A_{(i)})^2 \Big) \\
&\quad + \sqrt{x_{(n)} \cdot \mathfrak{m}(A_{(n)})} - x_{(n-1)} \cdot \mathfrak{m}(A_{(n)})^2 \Big\} \\
&= \mathfrak{C}_{\mathfrak{m}}^{(GM, F_{BPC})}(\vec{x}),
\end{aligned}
$$

since GM dominates F_{BPC} and GM is $(1, 0)$-increasing. Thus, $\mathfrak{C}_{\mathfrak{m}}^{(GM, F_{BPC})}$ is OD ($k, 0, \ldots, 0$)-increasing, for $k > 0$. $\qquad \square$

Theorem 2. *For any symmetric fuzzy measure* $\mathfrak{m} : 2^N \to [0, 1]$, $\mathfrak{C}_{\mathfrak{m}}^{(GM, F_{BPC})}$ *satisfies the boundary conditions* **(A2)**.

Proof. Consider $\vec{0} = (0, \ldots, 0) \in [0, 1]^n$ and $\vec{1} = (1, \ldots, 1) \in [0, 1]^n$. It follows that:

$$
\mathfrak{C}_{\mathfrak{m}}^{(GM, F_{BPC})}(\vec{0})
= \min \Big\{ 1, 0 + \sum_{i=2}^{n} \sqrt{0 \cdot \mathfrak{m}(A_{(i)})} - 0 \cdot \mathfrak{m}(A_{(i)})^2 \Big\} = 0
$$

$$
\mathfrak{C}_{\mathfrak{m}}^{(GM, F_{BPC})}(\vec{1})
= \min \{ 1, 1 + \sum_{i=2}^{n} \sqrt{1 \cdot \mathfrak{m}(A_{(i)})} - 1 \cdot \mathfrak{m}(A_{(i)})^2 \} = 1
$$

$\qquad \square$

Proposition 1. *For any symmetric fuzzy measure* $\mathfrak{m} : 2^N \to [0, 1]$, $\mathfrak{C}_{\mathfrak{m}}^{(GM, F_{BPC})}$ *is non-averaging*.

Proof. Suppose that $\mathfrak{C}_{\mathfrak{m}}^{(GM, F_{BPC})}$ is averaging. Now take $\vec{x} = (0.2, 0.5, 0.7, 0.9)$ and the power measure (Equation (10)), with $q = 1$. It follows that

$$
\begin{aligned}
&\mathfrak{C}_{\mathfrak{m}}^{(GM, F_{BPC})}(\vec{x}) \\
&= \min \Big\{ 1, 0.2 + \sum_{i=1}^{3} (GM(x_{(i)}, \mathfrak{m}(A_{(i)})) \\
&\qquad - F_{BPC}(x_{(i-1)}, \mathfrak{m}(A_{(i)}))) \Big\} \\
&= \min \Big\{ 1, 0.2 + \sum_{i=1}^{3} \Big(\sqrt{x_{(i)} \cdot \mathfrak{m}(A_{(i)})} - x_{(i-1)} \cdot \mathfrak{m}(A_{(i)})^2 \Big) \Big\} \\
&= \min \Big\{ 1, 0.2 + \sqrt{0.5 \cdot 0.75} - 0.2 \cdot 0.75^2 \\
&\qquad + \sqrt{0.7 \cdot 0.5} - 0.5 \cdot 0.5^2 + \sqrt{0.9 \cdot 0.25} - 0.7 \cdot 0.25^2 \Big\} \\
&= \min \{ 1, 1.59 \} > 0.9 = \max \{ 0.2, 0.5, 0.7, 0.9 \}.
\end{aligned}
$$

This a contradiction with the averaging property **(AV)**. $\qquad \square$

IV. USING $C_{F_1 F_2}$-INTEGRALS IN FUZZY RULE-BASED CLASSIFICATION SYSTEMS

In this section, our goal is to describe the main components of FRBCSs and the used fuzzy classifier. Furthermore, we present the considered FRM containing the main modification with respect to the original, which consist of the inclusion of the $C_{F_1 F_2}$-integrals in the aggregation stage.

A classification problem consists of t training examples $\mathrm{x}_p = (x_{p1}, \ldots, x_{pn}, y_p)$, with $p = 1, \ldots, t$, where x_{pi}, with $i = 1, \ldots, n$, is the value of the i-th variable and $y_p \in \mathbb{C} = \{C_1, \ldots, C_M\}$ is the label of the class of the p-th training example, and M is the number of classes.

In this paper, we focus on FRBCSs. Specifically, we use the Fuzzy Association Rule-based Classification model for High Dimensional Problems (FARC-HD [10]). The structure of the fuzzy rules generated by this classifier has the following form:

$$
\begin{aligned}
\text{Rule } R_j : \quad &\text{If } x_1 \text{ is } A_{j1} \text{ and } \ldots \text{ and } x_n \text{ is } A_{jn} \qquad (6) \\
&\text{then Class is } C_j \text{ with } RW_j,
\end{aligned}
$$

where R_j is the label of the j-th rule, A_{ji} is a fuzzy set representing a linguistic term modeled by a triangular shaped membership function. C_j is the class label and $RW_j \in [0, 1]$ is the rule weight [32], which in this case is computed as the confidence of the fuzzy rule.

In order to generate the set of fuzzy rules, FARC-HD applies the following three stages:

- *Fuzzy association rule extraction for classification*: In this step, an initial fuzzy rule base is obtained. To accomplish it, for each class, a search tree [33] is constructed, whose maximum depth is limited (parameter $depth_{max}$) . For each linguistic label (item), the support and confidence are calculated in order to obtain the frequent itemsets (set of items). Then, a fuzzy rule is generated for each frequent itemset.
- *Candidate rule prescreening*: This stage considers a weighting pattern scheme [34] to select the best generated fuzzy rules.
- *Genetic rule selection and lateral tuning*: At this point, the previously generated fuzzy rules are optimized so as

to enhance as much as possible the system's performance. To do so, the CHC evolutionary algorithm [35] is applied to carry out a rule selection process and the lateral tuning of the fuzzy sets [36].

A. Application of $C_{F_1F_2}$ in the fuzzy reasoning method

Once the knowledge base has been learnt and a new example has to be classified, the FRM is responsible to perform this task. As we have mentioned, we modify the classical FRM of FARC-HD [10] to include the usage of $C_{F_1F_2}$-integrals in its third stage. The steps of the new FRM are the following ones:

1) *Matching degree*: It represents the importance of the activation of the if-part of the rules for the example to be classified x_p, using a t-norm $T : [0,1]^2 \to [0,1]$.

$$\mu_{A_j}(x) = T(\mu_{A_{j1}}(x_1), \ldots, \mu_{A_{jn}}(x_n)), \quad j = 1, \ldots, L, \quad (7)$$

where L is the number of rules.

2) *Association degree*: For each rule, the matching degree is weighted by its rule weight:

$$b_j^k(x) = \mu_{A_j}(x) \cdot RW_j^k, \quad (8)$$

with $k = Class(R_j)$ and $j = 1, \ldots, L$.

3) *Example classification soundness degree for all classes*: This is the stage in which the $C_{F_1F_2}$-integrals are applied. At this point, for each class, all information given by the fired fuzzy rules is aggregated. To do so, the positive information provided by the previous step is aggregated by Equation (9).

$$S_k(x) = \mathfrak{C}_{\mathfrak{m}_k}^{C_{F_1F_2}}\left(b_1^k(x), \ldots, b_L^k(x)\right), \quad (9)$$

with k = Class(R_j) and $b_j^k > 0$, where $C_{F_1F_2}$ is the $C_{F_1F_2}$-integral considered to perform the aggregation. We remind that we use as F$_1$ and F$_2$ the functions presented Table III (Section III-B). $C_{F_1F_2}$-integrals are functions that generalize the Choquet integral (Equation (2)) and consequently, they use a fuzzy measure. In this work, we use the symmetric fuzzy measure applied in our previous papers, that is, the power measure:

$$\mathfrak{m}_k(X) = \left(\frac{|X|}{n}\right)^{q_k}, \text{ with } q_k > 0, \quad (10)$$

where the exponent q_k is genetically learnt (see section IV-B) by an evolutionary algorithm, to obtain the most suitable value, q_k, for each class k. Consequently, we use a different measure for each class.

4) *Classification*: The final decision is made in this step. To do so, a function $F : [0,1]^M \to \{1, \ldots, M\}$ is applied over the results obtained by example classification soundness degrees of all classes:

$$F((S_1, \ldots, S_M) = arg \max_{k=1,\ldots,M} (S_k). \quad (11)$$

B. Evolutionary learning of the fuzzy measure for each class

The original FARC-HD algorithm makes usage of the CHC evolutionary algorithm [35] to perform the lateral tuning of the fuzzy sets [36] and select the best set of fuzzy rules. In this paper we also learn a fuzzy measure for each class [4], k, by learning the q_k parameter as shown in Equation (10). The specific features of our evolutionary model are:

1) *Coding Scheme*: The chromosome is divided into three parts.
   (i) The first one considers the genes related to the tuning of lateral position of the membership functions and it has as many genes as the number of linguistic labels, where the range of each gene is [-0.5, 0.5] (for more details see [36]).
   (ii) The second part has one gene per class, k, and it is used to learn the exponent q_k. It is encoded in the range [0.01, 1.99]. However, as the real range is [0.01, 100] as showed in [4] the values of the genes have to be decoded in this range (See [4], [7] for details).
   (iii) The last part of the chromosome is related to the rule selection and it has as many genes as rules. Each gene determines if the corresponding rule is used in the FRM or not, by setting it to 1 (selected) or to 0 (not selected).

2) *Chromosome Evaluation*: We use as fitness function the standard accuracy rate, which is defined as follows.

$$Fitness(C) = \frac{\#Hits}{N},$$

where #Hits is the number of correctly classified examples and N is the total number of examples.

3) *Initial Gene Pool*: The population is composed by 50 individuals. Having one chromosome initialized by setting to 0 the value of all the genes to perform the lateral tuning, those used to learn the exponent of the fuzzy measure are set to 1.0 to obtain the classical cardinality fuzzy measure and the genes to perform the rule selection process are set to 1. The remainder chromosomes are randomly generated in the corresponding ranges of the genes.

4) *Crossover Operator*: We use the Parent Centric BLX (PCBLX) crossover operator [37] for the real coding part and the HUX [38] for the binary coded part. Two parents are crossed if their hamming distance divided by 2 is superior than the threshold Th, which is initialized as:

$$Th = \frac{(\#Genes \cdot BITSGENE)}{4.0} \quad (12)$$

We use the Gray code to convert each real coded gene to binary coding with a fixed number of bits for each gene (BITSGENE).

5) *Restarting Approach*: To increase the convergence of the algorithm, if new individuals are not included in the new population, we decrease the threshold by BITSGENE. When the threshold is smaller than 0 we pick the best chromosome (elitist scheme) and reset all the population with random values.

6) *Stopping Criteria*: The search process is stopped when:

(i) The maximum number of trials is reached.

(ii) A 100% is obtained as the fitness of the best individual.

V. EXPERIMENTAL FRAMEWORK

In this section we present the experimental framework used in this paper. We start by describing the datasets along with the configuration of the classifiers considered in this paper. After that, the statistical methods that are used for performance comparison.

A. Datasets and classifiers' set-up

In this study, to assess the performance of our approach, we consider 33 numeric datasets selected from the KEEL[2] dataset repository [39]. The features of the datasets are summarized in Table IV, showing for each one its identification (ID), followed by the name of the dataset (Dataset), the number of samples (#Samp.), the number of features (#Feat.) and the number of classes (#Class).

Examples containing missing information were removed, e.g., in the *wisconsin* dataset. Also, the datasets *magic*, *pageblocks*, *penbased*, *ring*, *satimage* and *twonorm* were stratified sampled at 10% in order to reduce their size for training.

For each dataset, we have considered a 5-fold cross-validation technique, that is, the dataset is split into five random partitions, with 20% of the examples and maintaining the class distribution. Then, we use four partitions for training, which involves the learning of the system and its optimization by applying the evolutionary algorithm, and the remainder is used for testing. This process is repeated five times, considering a different partition for testing each time. Moreover, we have repeated the process three times using a different seed in each one. Consequently, the result reported for each method is the average of the accuracy rate obtained in the fifteen testing folds.

In order to show the quality of our method, we compare it versus three state-of-the-art FRBCSs, namely, FURIA [11], IVTURS [9] and the original FARC-HD [10]. We show the configuration of these algorithms in Table V. In this table, we have to stress that our new proposal and IVTURS share the same fuzzy rule learning algorithm than that of FARC-HD and consequently, we use the same values for their parameters to perform a fair comparison.

B. Statistical tests for comparing performances

To give statistical support to the analysis of the results, we consider some hypothesis validation techniques [16], [40], that is, non-parametric tests, taking into account that the conditions that guarantee the reliance of the parametric tests cannot be warranted [17].

Specifically, we use the aligned Friedman rank test [41] to discover statistical differences among a group of results and to verify the quality of a method in comparison to other approaches. Observe that the algorithm achieving the lowest average ranking is the best one.

[2]http://www.keel.es

TABLE IV: Properties of the datasets considered in this study

Id.	Dataset	#Samp.	#Feat.	#Class
App	Appendicitis	106	7	2
Bal	Balance	625	4	3
Ban	Banana	5300	2	2
Bnd	Bands	365	19	2
Bup	Bupa	345	6	2
Cle	Cleveland	297	13	5
Con	Contraceptive	1473	9	3
Eco	Ecoli	336	7	8
Gla	Glass	214	9	6
Hab	Haberman	306	3	2
Hay	Hayes-Roth	160	4	3
Ion	Ionosphere	351	33	2
Iri	Iris	150	4	3
Led	led7digit	500	7	10
Mag	Magic	1,902	10	2
New	Newthyroid	215	5	3
Pag	Pageblocks	5,472	10	5
Pen	Penbased	10,992	16	10
Pho	Phoneme	5,404	5	2
Pim	Pima	768	8	2
Rin	Ring	740	20	2
Sah	Saheart	462	9	2
Sat	Satimage	6,435	36	7
Seg	Segment	2,310	19	7
Shu	Shuttle	58,000	9	7
Son	Sonar	208	60	2
Spe	Spectfheart	267	44	2
Tit	Titanic	2,201	3	2
Two	Twonorm	740	20	2
Veh	Vehicle	846	18	4
Win	Wine	178	13	3
Wis	Wisconsin	683	11	2
Yea	Yeast	1,484	8	10

TABLE V: Parameter setup of the considered algorithms

Algorithm	Configuration
FURIA	Number of optimizations: 2
	Number of folds: 3
FARC-HD, IVTURS and $C_{F_1F_2}$-integrals	Linguistic labels per variable: 5
	Conjunction operator: Product t-norm
	Rule weight: Confidence
	Minimum support: 0.05
	Minimum confidence: 0.8
	Depth of the search tree: 3
	Number of fuzzy rules that cover each example: 2
	Population size: 50
	Gray codification: 30 bits per gene
	Number of evaluations: 20.000

Moreover, we also use the Holm post-hoc test [42] to find the method that reject the equivalence hypothesis with respect to the best approach found with the aligned Friedman rank test. We compute the adjusted *p*-value (APV) considering that multiple tests are performed. Then, it is possible to directly compare the APV with the level of significance α, and, thus, we are capable of rejecting the null hypothesis.

Finally, we perform pair-wise comparisons by using the Wilcoxon test [43].

VI. EXPERIMENTAL RESULTS

This section is aimed at analyzing the performance of our new approach. To do so, we have separated the study in two parts. In the first one, we present the results obtained by the $C_{F_1F_2}$-integrals constructed using the pairs selected in section III-B (Table III). In the second one, in order to show the quality of our method, we perform comparisons against different state-of-the-art FRBCSs.

A. Analysis of the results of different $C_{F_1F_2}$-integrals

The results achieved in test by all the constructed $C_{F_1F_2}$-integrals are presented in Table VI. The rows represent the functions used as F_1, which are dominant in relation to the functions F_2, which are shown by columns and they have sub-ordination characteristics. The result of each cell is the average testing result and the standard deviation among the 33 datasets considered in the study. We have to point out that we only show the averaged results due to space limitations. The complete results can be accessed in – https://github.com/Giancarlo-Lucca/CF1F2-Integrals. In Table VI we highlight in **boldface** the maximum accuracy per row and we underline the best accuracy for each column. Observe that blank spaces are related to combinations that could not be performed, since the dominance property is not satisfied for the specific pair of functions.

In a general looking, it is possible to observe that the largest accuracy is obtained by picking the function F_{GL} as F_1 and T_M as function F_2. This pair is a combination of a function having a high dominance as F_1 combined with a function with a low subordination as F_2. Moreover, we can observe that for the functions to be F_1 the results are better when they are paired with a function F_2 with a high subordination degree (results highlighted in **boldface**). The opposite is also observed, since for each F_2 function, the best results are obtained when considering a F_1 with high dominance (underlined results).

Analyzing the results by categories (high, medium and low) according to the functions F_1, we have that:

- Using a function with high dominance characteristics as F_1 provides good results, since eight of the top ten best classifications, are pairs with this characteristic. Observe that, if we pick the functions GM and F_{GL} as F_1, the results tend to present a stability since the accuracies could be considered as similar. Regarding the sine function, S, its unsatisfactory behavior could occur since the differences between the pair of functions are too wide, which may imply a decrease on the performance of the classifier. Observe that in [30] this function also presents a similar behavior.
- The usage of functions having medium dominance characteristic as F_1 (T_{HP}, T_M and F_{IM}) produced 16 possible combinations. The combination of these functions with functions having a low subordination (T_M) achieved an accuracy mean inferior than 80%. The same happened with the medium subordination function, F_{NA} and T_{HP}. We believe that this occurred because these functions produce generally averaging combinations and, according to our previous paper [12], the results when using non-averaging operators excels those of averaging ones. For the remaining cases, the achieved mean is superior than 80%.
- Applying functions with low dominance as F_1, in general, does not fulfill the dominance property and, for this reason, less pairs can be used to construct $C_{F_1F_2}$-integrals. However, from the seven pairs constructed in this study, three of them provide poor results (less than 80%) and

the remainder ones obtain satisfactory results.

B. Comparisons against other non-averaging aggregation functions and state-of-the-art fuzzy classifiers

As mentioned before, in general the obtained results tend to be stable and satisfactory. Thus, in order to demonstrate the quality of our approach, we compared the performance of the $C_{F_1F_2}$-integral that achieved the highest accuracy ($F_{GL}-T_M$) against the best non-averaging function of our previous paper [12] (F_{NA2}), a classical non-averanging aggregation operator like the probabilistic sum (P^*) and three state-of-the-art fuzzy classifiers, namely, FURIA [11], IVTURS [9] and FARC-HD [10].

The results achieved in testing by the different methods, are detailed in Table VII by columns. In each row of this table we present the accuracy and the standard deviation obtained per each dataset. Furthermore, we highlight in **boldface** the best achieved result for each one and, in the two last rows, we present the number of datasets in which the classifier achieves the best (#Wins) and the worst result (#Losses).

From the obtained results, performing just a simple numerical comparison, it is possible to observe that FURIA is the method achieving the best global mean and the largest number of best classification results. The $C_{F_1F_2}$-integral achieves the second position in both criteria. However, we have to stress that whilst FURIA provides the worst result in fourteen datasets, while our approach achieves the worst results in a single case. Therefore, we can observe that our method provides a good performance in a regular way. This affirmation can also be made to the CF-integral, F_{NA2}, but in this case it provides the best results in a less number of datasets. For the remainder methods, the results are worse than those of FURIA and our new approach, since the number of datasets having the best results are less and the number of loss cases are larger.

In order to highlight the behavior or our new method, if we look at the results in Table VI and we compare them against the ones of Table VII, we can find a large number of combinations leading to a global mean equal or larger than that of the compared methods (except that of FURIA). Specifically, the number of combinations having an equal or greater average result is 43, 32, 29 and 28 when compared against IVTURS, P^*, FARC-HD and F_{NA2}, respectively.

To support these findings we have conducted a set of statistical studies (as many as combinations in Table VII) using the aligned Friedman rank test to compare each $C_{F_1F_2}$-integral with the remainder methods considered in this section, whose obtained results are available in Table VIII. Specifically, in this table we only show the results of those $C_{F_1F_2}$-integrals (in columns) that obtain the best rank and consequently, they are used as control method in the post-hoc Holm's test, whose obtained APV is shown in brackets. If there are statistical differences between two methods, considering 0.10 (10%) as the level of confidence, we underline the APV.

From the obtained results, it is noticeable that nine different $C_{F_1F_2}$-integrals are considered as control method, presenting statistical differences against IVTURS (all cases), FARC-HD

TABLE VI: Accuracy mean achieved in testing by different $C_{F_1 F_2}$-integrals

| | | | F_2 (Subordination) | | | | | | | |
| | | | | Low | | | Medium | | | High | |
		S	GM	T_M	T_P	F_{NA}	T_{HP}	T_{DP}	F_{BPC}	$T_Ł$
Low	T_{DP}					76.74 ± 3.54		**76.96** ± 3.54		
	F_{NA}							**80.41** ± 3.18		
	O_B				79.96 ± 3.14			<u>80.46</u> ± 3.08	80.18 ± 3.14	80.43 ± 3.31
	T_{HP}				80.18 ± 3.34		78.64 ± 3.79	80.40 ± 3.16	80.37 ± 3.26	80.35 ± 3.33
Medium	T_M			79.94 ± 3.25	80.30 ± 3.16	79.43 ± 3.45	79.90 ± 3.22	80.45 ± 3.19	80.49 ± 3.24	**80.53** ± 3.10
	F_{IM}			79.89 ± 3.24	80.26 ± 3.28	79.95 ± 3.23	80.18 ± 3.21	80.26 ± 3.05	80.04 ± 3.18	**80.27** ± 3.21
	GM		79.72 ± 3.12	80.49 ± 3.27	80.46 ± 3.34	80.43 ± 3.36	<u>80.50</u> ± 3.24	80.44 ± 3.39	**80.68** ±0.77	<u>80.64</u> ± 3.24
High	F_{GL}		<u>80.48</u> ± 3.21	**80.72** ± 3.33	<u>80.60</u> ± 3.13	<u>80.54</u> ± 3.31	80.43 ± 3.21	80.40 ± 3.18	80.65 ± 3.25	80.41 ± 3.29
	S	<u>76.90</u> ± 3.56	79.92 ± 3.31	79.70 ± 3.29	79.89 ± 3.18	79.65 ± 3.40	79.75 ± 3.34	79.97 ± 3.37	80.00 ± 3.22	**80.21** ±3.22

TABLE VII: Results achieved in testing by different FRMs

Dataset	FURIA	IVTURS	FARC-HD	P*	F_{NA2}	(F_{GL}, T_M)
App	**87.45** ± 4.46	83.97 ± 5.03	83.65 ± 6.92	84.60 ± 5.75	85.83 ± 4.34	83.64 ± 7.37
Bal	83.57 ± 1.45	85.28 ± 1.70	86.83 ± 1.74	86.56 ± 2.15	87.84 ± 1.94	**87.89** ± 2.02
Ban	**88.43** ± 1.10	81.55 ± 1.20	85.84 ± 1.34	85.18 ± 1.10	84.58 ± 1.20	84.66 ± 0.89
Bnd	65.42 ± 5.53	67.09 ± 3.82	69.11 ± 5.51	69.31 ± 6.83	69.58 ± 5.26	**70.35** ± 7.43
Bup	66.67 ± 3.97	64.73 ± 4.46	66.09 ± 3.84	62.61 ± 5.78	63.38 ± 5.07	**67.15** ± 5.03
Cle	56.35 ± 2.70	58.81 ± 4.31	56.55 ± 3.82	**59.03** ± 2.60	56.21 ± 5.78	57.57 ± 3.93
Con	**54.74** ± 1.45	53.16 ± 1.94	53.23 ± 2.04	53.16 ± 2.50	53.75 ± 2.17	53.54 ± 1.80
Eco	80.26 ± 3.89	79.27 ± 3.40	81.35 ± 3.59	80.66 ± 3.66	80.37 ± 4.65	**81.55** ± 3.94
Gla	**71.35** ± 6.63	66.69 ± 8.20	65.44 ± 6.83	65.75 ± 5.23	66.99 ± 5.60	65.91 ± 5.42
Hab	72.54 ± 3.14	72.63 ± 8.27	**72.75** ± 5.46	71.00 ± 4.56	71.11 ± 6.31	71.22 ± 6.08
Hay	**81.00** ± 7.21	80.51 ± 6.24	78.21 ± 7.16	78.44 ± 7.65	78.44 ± 7.67	79.97 ± 6.02
Ion	90.71 ± 2.76	**91.66** ± 3.52	89.37 ± 4.55	87.85 ± 5.42	88.99 ± 3.63	90.70 ± 4.21
Iri	94.22 ± 5.24	**96.22** ± 3.95	94.44 ± 3.98	94.67 ± 4.21	94.67 ± 4.21	94.89 ± 4.64
Led	**71.40** ± 6.27	69.73 ± 5.49	69.73 ± 5.75	69.00 ± 5.65	69.87 ± 5.41	69.67 ± 5.71
Mag	**80.63** ± 1.82	79.90 ± 2.76	80.39 ± 2.48	80.13 ± 2.13	79.32 ± 2.37	80.34 ± 2.44
New	94.57 ± 2.91	95.50 ± 2.52	95.19 ± 3.17	94.88 ± 3.14	95.04 ± 2.64	**96.12** ± 2.82
Pag	**95.50** ± 0.95	94.82 ± 1.66	94.40 ± 1.33	94.34 ± 1.28	94.82 ± 1.77	94.64 ± 1.63
Pen	91.67 ± 2.20	92.18 ± 2.94	**92.58** ± 2.27	92.24 ± 2.00	92.06 ± 2.86	92.30 ± 2.38
Pho	**84.91** ± 1.10	80.13 ± 0.99	81.85 ± 1.06	81.78 ± 1.42	81.21 ± 1.49	81.40 ± 0.76
Pim	**75.39** ± 2.57	74.78 ± 2.60	74.65 ± 2.20	74.74 ± 2.72	75.04 ± 2.32	74.87 ± 1.72
Rin	85.68 ± 3.89	88.24 ± 2.80	**91.08** ± 2.03	90.54 ± 2.41	89.46 ± 2.63	90.95 ± 1.69
Sah	70.13 ± 3.35	**71.28** ± 3.04	69.33 ± 2.67	69.98 ± 4.65	69.62 ± 2.69	70.84 ± 4.67
Sat	**81.91** ± 1.62	75.95 ± 2.16	79.83 ± 1.81	80.14 ± 2.43	80.25 ± 2.29	79.78 ± 2.33
Seg	**97.20** ± 0.91	90.53 ± 1.55	93.03 ± 1.63	92.86 ± 1.24	92.55 ± 1.64	93.48 ± 1.44
Shu	**99.68** ± 0.13	91.34 ± 2.01	94.51 ± 2.53	95.00 ± 2.40	97.12 ± 0.91	96.21 ± 0.96
Son	79.55 ± 7.02	79.36 ± 4.05	79.99 ± 4.88	81.29 ± 4.87	82.24 ± 4.79	**82.43** ± 6.60
Spe	78.00 ± 8.13	**80.28** ± 3.35	78.51 ± 4.03	78.63 ± 4.55	79.39 ± 2.64	78.90 ± 2.90
Tit	78.30 ± 1.34	**78.87** ± 1.48	**78.87** ± 1.48	**78.87** ± 1.48	**78.87** ± 1.48	**78.87** ± 1.48
Two	87.75 ± 3.20	**93.06** ± 1.37	90.68 ± 2.35	89.91 ± 2.00	91.89 ± 2.01	92.21 ± 1.75
Veh	**70.45** ± 3.33	66.00 ± 2.36	69.03 ± 2.85	68.84 ± 3.33	67.65 ± 2.23	69.94 ± 3.08
Win	94.15 ± 3.79	95.50 ± 2.78	95.48 ± 4.37	95.67 ± 3.88	95.50 ± 4.59	**96.82** ± 4.53
Wis	96.49 ± 1.14	96.54 ± 1.16	96.58 ± 1.06	96.73 ± 1.02	**96.93** ± 1.02	96.78 ± 0.89
Yea	58.09 ± 1.87	55.77 ± 2.47	**58.42** ± 1.29	58.29 ±2.20	56.94 ± 1.51	58.22 ± 1.44
Mean	**80.73** ± 3.24	79.74 ± 3.20	80.21 ± 3.27	80.08 ± 3.40	80.23 ± 3.25	80.72 ± 3.33
#Wins	14	6	5	2	2	8
#Losses	10	10	4	5	2	1

(twice), P* (four times) and $C_{F_1 F_2}$(two times). Regarding FURIA, we can infer that there are no statistical differences between both methods, since the obtained APVs are high.

In order to complete the statistical study we have performed a set of pairwise comparisons using the Wilcoxon's statistical test. Specifically, we compare the nine different $C_{F_1 F_2}$-integrals, shown in Table VIII, against the same methods in that table. The results are available in Table IX, in which we present for each comparison the p-value and between parenthesis the rank obtained by the $C_{F_1 F_2}$-integral. We must stress that as the total sum of ranks of this test is 561, the rank of the compared method can be calculated by the difference. Consequently, if the rank, shown in Table IX, is superior than 280.5 it means that is favorable to our approach.

The results in Table IX reinforce that the considered $C_{F_1 F_2}$-integrals are equivalent to FURIA. However, looking at these results we can observe that all the $C_{F_1 F_2}$-integrals are statistically improving the remainder methods in all cases. Therefore, we can conclude that our new methodology presents a competitive performance versus state-of-the-art fuzzy classifiers.

VII. CONCLUSION

In this paper we have defined the concept of $C_{F_1 F_2}$-integrals, which are a generalization of the CC-integral introduced in [7]. Specifically, these integrals use two different fusion functions, F_1 and F_2, in order to try to enhance the behavior of FRBCSs. The constructed $C_{F_1 F_2}$-integrals are non-averaging as most of the aggregation operators used by state-of-the-art fuzzy classifiers are. Furthermore, these integrals are OD increasing functions satisfying proper boundary conditions. We have presented a method to select the best combination of functions to be F_1 and F_2, which is based on the concept of dominance and subordination.

From the obtained results, we can conclude that the results of this approach could be considered as satisfactory and stable, since the results are quite similar in many cases. Furthermore, we showed that different $C_{F_1 F_2}$-integrals provide competitive results when compared against FURIA. Moreover, we have to highlight that we enhance two state-of-the-art fuzzy classifiers like IVTURS and FARC-HD and two non-averaging operators like F_{NA2} and the probabilistic sum. All these facts support that this approach is an efficient option and it expands the scope of the generalizations of the Choquet integral.

REFERENCES

[1] R. O. Duda, P. E. Hart, D. G. Stork, Pattern Classification (2Nd Edition), Wiley-Interscience, 2000.

[2] H. Ishibuchi, T. Nakashima, M. Nii, Classification and Modeling with Linguistic Information Granules, Advanced Approaches to Linguistic Data Mining, Advanced Information Processing, Springer, Berlin, 2005.

[3] 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.

[4] E. Barrenechea, H. Bustince, J. Fernandez, D. Paternain, J. A. Sanz, Using the Choquet integral in the fuzzy reasoning method of fuzzy rule-based classification systems, Axioms 2 (2) (2013) 208–223.

[5] G. Choquet, Theory of capacities, Annales de l'Institut Fourier 5 (1953–1954) 131–295.

[6] G. Lucca, J. Sanz, G. Pereira Dimuro, B. Bedregal, R. Mesiar, A. Kolesárová, H. Bustince Sola, Pre-aggregation functions: construction and an application, IEEE Transactions on Fuzzy Systems 24 (2) (2016) 260–272.

[7] 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.

[8] G. Lucca, J. A. Sanz, G. P. Dimuro, B. Bedregal, J. Fernández, H. Bustince, Analyzing the behavior of a CC-integral in a fuzzy rule-based classification system, in: 2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), 2017, pp. 1–6.

TABLE VIII: Statistical results using the align Friedman rank test and Holm's post hoc test

	F_{GL}–T_M	GM–F_{BPC}	F_{GL}–F_{BPC}	GM–$T_Ł$	F_{GL}–T_P	F_{GL}–F_{NA}	T_M–$T_Ł$	GM–T_{HP}	GM–T_M
$C_{F_1F_2}$	74.06 (-)	74.75 (-)	76.68 (-)	75.01 (-)	81.65 (-)	81.63 (-)	82.84 (-)	81.09 (-)	81.22 (-)
FURIA	86.45 (0.37)	86.03 (0.42)	85.98 (0.50)	85.93 (0.43)	85.31 (0.79)	85.62 (0.77)	84.46 (0.90)	84.51 (0.80)	84.54 (0.81)
F_{NA2}	104.15 (0.09)	104.84 (0.09)	104.72 (0.14)	103.75 (0.12)	103.71 (0.35)	102.78 (0.40)	102.56 (0.48)	103.28 (0.34)	103.18 (0.35)
FARC-HD	104.63 (0.09)	102.84 (0.09)	102.34 (0.14)	103.92 (0.12)	101.6 (0.35)	101.95 (0.40)	101.33 (0.48)	101.28 (0.34)	102.18 (0.35)
P*	109.69 (0.04)	109.65 (0.05)	109.21 (0.08)	110.6 (0.04)	108.53 (0.22)	109.36 (0.19)	109.51 (0.23)	110.06 (0.16)	109.51 (0.17)
IVTURS	117.99 (0.00)	118.86 (0.00)	118.04 (0.01)	117.75 (0.01)	116.18 (0.07)	115.63 (0.07)	116.27 (0.08)	116.75 (0.05)	116.34 (0.06)

TABLE IX: Statistical results using the Wilcoxon's test

	F_{GL}–T_M	GM–F_{BPC}	F_{GL}–F_{BPC}	GM–$T_Ł$	F_{GL}–T_P	F_{GL}–F_{NA}	T_M–$T_Ł$	GM–T_{HP}	GM–T_M
FURIA	0.99 (281)	0.87 (272)	0.85 (270)	0.72 (261)	0.71 (260)	0.66 (256)	0.41 (235)	0.45 (239)	0.54 (247)
F_{NA2}	0.01 (420)	0.01 (420)	0.03 (398.5)	0.01 (419.5)	0.05 (385.5)	0.03 (397.5)	0.10 (369.5)	0.05 (385)	0.08 (377)
IVTURS	0.02 (408)	0.01 (419.5)	0.01 (420.5)	0.02 (410.5)	0.04 (391.5)	0.05 (388.5)	0.08 (369)	0.06 (383)	0.07 (379)
P*	0.00 (465.5)	0.00 (451)	0.00 (452.5)	0.00 (465.5)	0.00 (443)	0.00 (430.5)	0.00 (421.5)	0.00 (462.5)	0.00 (454.5)
FARC-HD	0.00 (451.5)	0.01 (421)	60.4E-0.4 (472.5)	0.00 (442.5)	0.00 (429)	0.05 (387.5)	0.08 (377.5)	0.05 (388)	0.01 (415)

[9] 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.

[10] 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.

[11] J. Hühn, E. Hüllermeier, Furia: an algorithm for unordered fuzzy rule induction, Data Mining and Knowledge Discovery 19 (3).

[12] 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.

[13] J. Alcalá-Fdez, L. Sánchez, S. García, M. Jesus, S. Ventura, J. Garrell, J. Otero, C. Romero, J. Bacardit, V. Rivas, J. Fernández, F. Herrera, Keel: a software tool to assess evolutionary algorithms for data mining problems, Soft Computing 13 (3) (2009) 307–318.

[14] C. Alsina, M. J. Frank, B. Schweizer, Associative Functions: Triangular Norms and Copulas, World Scientific Publishing Company, Singapore, 2006.

[15] J. Derrac, S. García, D. Molina, F. Herrera, A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms, Swarm and Evolutionary Computation 1 (1) (2011) 3 – 18.

[16] S. García, A. Fernández, J. Luengo, F. Herrera, A study of statistical techniques and performance measures for genetics–based machine learning: Accuracy and interpretability, Soft Computing 13 (10) (2009) 959–977.

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

[18] G. Beliakov, A. Pradera, T. Calvo, Aggregation Functions: A Guide for Practitioners, Springer, Berlin, 2007.

[19] G. Mayor, E. Trillas, On the representation of some aggregation functions, in: Proceedings of IEEE International Symposium on Multiple-Valued Logic, IEEE, Los Alamitos, 1986, pp. 111–114.

[20] H. Bustince, J. Fernandez, A. Kolesárová, R. Mesiar, Directional monotonicity of fusion functions, European Journal of Operational Research 244 (1) (2015) 300–308.

[21] T. Murofushi, M. Sugeno, M. Machida, Non-monotonic fuzzy measures and the Choquet integral, Fuzzy Sets and Systems 64 (1) (1994) 73 – 86.

[22] G. Beliakov, H. Bustince, T. Calvo, A Practical Guide to Averaging Functions, Springer, Berlin, New York, 2016.

[23] E. P. Klement, R. Mesiar, E. Pap, Triangular Norms, Kluwer Academic Publisher, Dordrecht, 2000.

[24] H. Bustince, J. Fernandez, R. Mesiar, J. Montero, R. Orduna, Overlap functions, Nonlinear Analysis: Theory, Methods & Applications 72 (3-4) (2010) 1488–1499.

[25] R. B. Nelsen, An introduction to copulas, Vol. 139 of Lecture Notes in Statistics, Springer, New York, 1999.

[26] G. P. Dimuro, B. Bedregal, Archimedean overlap functions: The ordinal sum and the cancellation, idempotency and limiting properties, Fuzzy Sets and Systems 252 (2014) 39 – 54.

[27] G. P. Dimuro, B. Bedregal, Additive generators of overlap functions, in: H. Bustince, J. Fernandez, R. Mesiar, T. Calvo (Eds.), Aggregation Functions in Theory and in Practice, Vol. 228 of Advances in Intelligent Systems and Computing, Springer, Berlin, 2013, pp. 167–178.

[28] G. P. Dimuro, B. Bedregal, H. Bustince, M. J. Asiáin, R. Mesiar, On additive generators of overlap functions, Fuzzy Sets and Systems 287 (2016) 76 – 96, theme: Aggregation Operations.

[29] G. P. Dimuro, B. Bedregal, On residual implications derived from overlap functions, Information Sciences 312 (2015) 78 – 88.

[30] M. Elkano, M. Galar, J. Sanz, A. Fernández, E. Barrenechea, F. Herrera, H. Bustince, Enhancing multi-class classification in FARC-HD fuzzy classifier: On the synergy between n-dimensional overlap functions and decomposition strategies, IEEE Transactions on Fuzzy Systems 23 (5) (2015) 1562–1580.

[31] H. Bustince, E. Barrenechea, M. Sesma-Sara, J. Lafuente, G. P. Dimuro, R. Mesiar, A. Kolesárová, Ordered directionally monotone functions. justification and application, IEEE Transactions on Fuzzy Systems (2017) 1(In press, corrected proof). doi:10.1109/TFUZZ.2017.2769486.

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

[33] R. Agrawal, R. Srikant, Fast algorithms for mining association rules in large databases, in: Proceedings of the 20th International Conference on Very Large Data Bases, VLDB '94, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1994, pp. 487–499.

[34] B. Kavšek, N. Lavrač, V. Jovanoski, Apriori-sd: Adapting association rule learning to subgroup discovery, in: M. R. Berthold, H.-J. Lenz, E. Bradley, R. Kruse, C. Borgelt (Eds.), Advances in Intelligent Data Analysis V: 5th International Symposium on Intelligent Data Analysis, IDA 2003, Berlin, Germany, August 28-30, 2003. Proceedings, Springer Berlin Heidelberg, Berlin, Heidelberg, 2003, pp. 230–241.

[35] L. J. Eshelman, The CHC adaptive search algorithm: How to have safe search when engaging in nontraditional genetic recombination, in: G. J. E. Rawlings (Ed.), Foundations of Genetic Algorithms, Morgan Kaufmann, San Francisco, 1991, pp. 265–283.

[36] 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.

[37] 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.

[38] L. J. Eshelman, J. D. Schaffer, Real-coded genetic algorithms and interval-schemata, in: L. D. WHITLEY (Ed.), Foundations of Genetic Algorithms, Vol. 2 of Foundations of Genetic Algorithms, Elsevier, 1993, pp. 187 – 202.

[39] J. Alcalá-Fdez, L. Sánchez, S. García, M. Jesus, S. Ventura, J. Garrell, J. Otero, C. Romero, J. Bacardit, V. Rivas, J. Fernández, F. Herrera, Keel: a software tool to assess evolutionary algorithms for data mining problems, Soft Computing 13 (3) (2009) 307–318.

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

[41] 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.

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

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

Giancarlo Lucca has obtained his B.Sc. in information sciences and M.Sc. degrees in computer engineering, both at Universidade Federal do Rio Grande, Brazil. He obtained his Ph.D at Universidad Publica de Navarra, Spain, under the advising of Prof. Humberto Bustince and Prof. José A. Sanz.

José Antonio Sanz received the M.Sc. degree in computer sciences and the Ph.D. degree, both form the Public University of Navarre. He is currently an Associate Lecturer with the Department of Automatics and Computation, Public University of Navarre. He is the author of 27 published original articles in international journals. He received the best paper award in the FLINS 2012 international conference and the Pepe Millá award in 2014.

Graçaliz Pereira Dimuro received the M.Sc. and Ph.D. degrees from the Instituto de Informática of Universidade Federal do Rio Grande do Sul, Brazil. In 2015, she had a Pos-Doc post-doctorate grant from the Science Without Borders Program from the Brazilian Research Funding Agency CNPq, to join GIARA research group at Universidad Publica de Navarra, and, in 2017, she had a talent grant at the Institute of Smart Cities of Universidad Publica de Navarra, Spain. Currently, she is a full professor with Universidade Federal do Rio Grande, Brazil, and a Researcher of CNPq, Brazil.

Javier Fernández received the M.Sc. and Ph.D. degrees in mathematics from the University of Zaragoza, Spain. He is currently an Associate Lecturer with the Department of Automatics and Computation, Public University of Navarre, Spain. He is involved with teaching artificial intelligence and computational mathematics for students of the computer sciences.

Humberto Bustince is full professor in the Public University of Navarra and Honorary Professor in the University of Nottingham. He has authored more than 210 works. He is associated editor of the IEEE Transactions on Fuzzy Systems journal and member of the editorial board of the journals Fuzzy Sets and Systems, Information Fusion, International Journal of Computational Intelligence Systems and Journal of Intelligent & Fuzzy Systems.

Benjamín Bedregal received the Engineering degree in Computing from the Tarapac University (UTA), in Arica, Chile, the M.Sc. degree in informatics and the Ph.D. degree in computer sciences from the Federal University of Pernambuco, Brazil, in 1987 and 1996, respectively. He is a titular Professor at the Department of Informatics and Applied Mathematics, Federal University of Rio Grande do Norte, Natal, Brazil.