Information Sciences 435 (2018) 94–110

Contents lists available at ScienceDirect

Information Sciences

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

# $C_F$-integrals: A new family of pre-aggregation functions with application to fuzzy rule-based classification systems

Giancarlo Lucca[a,*], José Antonio Sanz[a,b], Graçaliz Pereira Dimuro[b,c], Benjamín Bedregal[d], Humberto Bustince[a,b], Radko Mesiar[e,f]

[a] Depto. of Automática y Computación, Universidad Publica de Navarra, Campus Arrosadía s/n, P.O. Box 31006, Pamplona, Spain
[b] Institute of Smart Cities, Universidad Publica de Navarra, Campus Arrosadía, Navarra, 31006, Spain
[c] Centro de Ciências Computacionais, Universidade Federal do Rio Grande, Av. Itália km 08, Campus Carreiros, Rio Grande, 96201-900, Brazil
[d] Departamento de Informática e Matemática Aplicada, Universidade Federal do Rio Grande do Norte, Campus Universitário, Natal, 59072-970, Brazil
[e] Department of Mathematics, Slovak University of Technology, Radlinského 11, Bratislava 81005, Slovakia
[f] Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Prague 18208, Czech Republic

A R T I C L E I N F O

*Article history:*
Received 28 April 2017
Revised 29 September 2017
Accepted 23 December 2017
Available online 24 December 2017

*Keywords:*
$C_F$-integral
Choquet integral
Pre-aggregation function
Classification problems
Fuzzy reasoning method
Fuzzy rule-based classification systems

A B S T R A C T

This paper introduces the family of $C_F$-integrals, which are pre-aggregations functions that generalizes the Choquet integral considering a bivariate function $F$ that is left 0-absorbent. We show that $C_F$-integrals are $\vec{1}$-pre-aggregation functions, studying in which conditions they are idempotent and/or averaging functions. This characterization is an important issue of our approach, since we apply these functions in the Fuzzy Reasoning Method (FRM) of a fuzzy rule-based classification system and, in the literature, it is possible to observe that non-averaging aggregation functions usually provide better results. We carry out a study with several subfamilies of $C_F$-integrals having averaging or non-averaging characteristics. As expected, the proposed non-averaging $C_F$-integrals obtain more accurate results than the averaging ones, thus, offering new possibilities for aggregating accurately the information in the FRM. Furthermore, it allows us to enhance the results of classical FRMs like the winning rule and the additive combination.



## 1. Introduction

An effective approach to handle classification problems [25] is through the application of the Fuzzy Rule-Based Classification Systems (FRBCSs) [35], since they provide the user with interpretable models by using linguist labels in their rules and, moreover, achieving accurate results. FRBCSs have been applied in several problems, including real-time vehicle classification [57], health [52] or economy [49], among many others.

A key component in any FRBCS is the Fuzzy Reasoning Method (FRM) [14], which determines how the information learned in form of fuzzy rules will be used to classify new examples. A crucial point in any FRM is the way to obtain the

* Corresponding author.
*E-mail addresses:* lucca.112793@e.unavarra.es (G. Lucca), joseantonio.sanz@unavarra.es (J. Antonio Sanz), gracalizdimuro@furg.br, gracaliz.pereira@unavarra.es (G.P. Dimuro), bedregal@dimap.ufrn.br (B. Bedregal), bustince@unavarra.es (H. Bustince), mesiar@math.sk (R. Mesiar).

https://doi.org/10.1016/j.ins.2017.12.029
0020-0255/

information associated with each class of the problem. This is done by applying an aggregation [8,30,44] or, more recently, a pre-aggregation [11,12,22,40,42] function over the local information given by each fired rule of the FRBCS.

In the literature, it is possible to find classical FRMs that consider, as the aggregation operator, the maximum (Winning Rule – WR) or the normalized sum (Additive Combination – AC). The first method takes into consideration only one rule (the one having the maximum compatibility with the example to be classified) and, obviously, has an averaging and idempotent behavior. On the other hand, the second method aggregates the information of all triggered rules and it has neither an averaging nor an idempotent behavior. Usually, the FRM of AC provides better performance than that the FRM of the WR, as it can be observed widely in the literature, since the most accurate FRBCSs currently (FURIA [33], IVTURS [51] and FARC-HD [2]) make use of the AC.

Recently, several works were proposed to apply aggregation and pre-aggregation functions (with averaging and idempotent characteristics) to aggregate the local information associated with each rule. The initial idea was proposed by Barrenechea et al. [4], where the Choquet integral [13] was used to perform this aggregation in a way that also took into account the correlation between the rules. After that, this method was improved by Lucca et al. [40], introducing the concept of pre-aggregation function, which is a generalization of the Choquet integral where the product operator of this function is replaced by a t-norm [36]. In [41], the Choquet integral in its expanded form was generalized using copula functions [3], instead of the product operator,obtaining aggregation functions called CC-integrals.

In this paper, the product operator of the Choquet integral is replaced by a more general function $F$: $[0, 1]^2 \to [0, 1]$. We study which are the minimal requirements that this function $F$ must satisfy so that the obtained generalization of the Choquet integral is a pre-aggregation function. Specifically, we have found that the key property to achieve this is the presence of 0 as a left annihilator element, in which case the function $F$ is called left 0-absorbent.

The general aim is to apply such pre-aggregation functions in the FRM of a FRBCS, searching for more flexible ways of aggregating information. In this manner, it is possible to make an in-depth analysis of the their performances according to their averaging or non-averaging behavior. Observe that the non-averaging behavior is a novel approach, since we have not considered it in our previous works.

Then, the first objective of this paper is the definition of the concept of $C_F$-integral, which is a generalization of the Choquet integral based on a left 0-absorbent function $F$ satisfying a minimal set of properties that guarantees that any $C_F$-integral is a pre-aggregation function. Secondly, we analyze under which conditions such $C_F$-integrals are idempotent and/or averaging pre-aggregation functions. In the sequence, we study subfamilies of $C_F$-integrals, considering left 0-absorbent functions $F$ that are (I) t-norms [36], (II) overlap functions [5,10,19,20,23,24], (III) copulas [3] that are neither t-norms nor overlap functions, (IV) other kinds of aggregation functions and (V) pre-aggregation functions.

As done in [4,40,41], we apply this generalization in the FRM of FRBCSs and we conduct an experimental study composed of two steps. The first one is based on $C_F$-integrals having averaging characteristics, where we compare them among themselves in order to choose the representative for this family. After that, we compare this representative against the classical FRM of WR, the standard Choquet integral, the best pre-aggregation achieved in [40] and the best CC-integral obtained in [41].

The second part of this analysis is concerned with $C_F$-integrals having non-averaging characteristics. As done in the first part of the experimental study, firstly we determine the best function of this family and compare it against the classical non-averaging FRMs of AC and probabilistic sum.

The experimental study was performed considering 33 datasets that are available in the KEEL database repository [1]. The standard accuracy rate is used to measure the performance of the classifiers and the results are supported by appropriate statistical tests [15,28,53].

The paper is organized in the following way. Section 2 is aimed at introducing the basic concepts that are necessary to understand the paper. The concept of $C_F$-integral is introduced in Section 3, where we analyze several properties, such as idempotency and averaging behaviors. The Section 4 presents the methodology to build a generalized FRM of FRBCSs using different $C_F$ integrals, configurations of the classifier used in this paper and the experimental framework. In Section 5 we show the experimental study, showing the results achieved in test considering this new approach, and the appropriate analysis. The conclusions are drawn in Section 6.

## 2. Basic concepts

This section presents the preliminary concepts that are used in the development of this work. In our approach, the basic property that is considered for any bivariate function defined on [0, 1], is the following.

**Definition 1.** A bivariate function $F$: $[0, 1]^2 \to [0, 1]$ with 0 as left annihilator element, that is, satisfying:

(**LAE**) $\forall y \in [0, 1] : F(0, y) = 0,$

is said to be left 0-absorbent.

Moreover, the following two basic properties are also important:

(**RNE**) Right Neutral Element: $\forall x \in [0, 1] : F(x, 1) = x;$
(**LC**) Left Conjunctive Property: $\forall x, y \in [0, 1]$: $F(x, y) \leq x$;

Any bivariate function $F$: $[0, 1]^2 \to [0, 1]$ satisfying both **(LAE)** and **(RNE)** is called left 0-absorbent **(RNE)**-function.

Now, we recall the concepts of aggregation and pre-aggregation functions, and specific types of aggregation functions, such as t-norms, overlap and copulas.

**Definition 2** [8,30,44]**.** A function $A$: $[0, 1]^n \to [0, 1]$ is an aggregation function if the following conditions hold:

(**A1**) $A$ is increasing[1] in each argument: for each $i \in \{1, \ldots, n\}$, if $x_i \le y$, then $A(x_1, \ldots, x_n) \le A(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n)$;
(**A2**) $A$ satisfies the boundary conditions: (i) $A(0, \ldots, 0) = 0$ and (ii) $A(1, \ldots, 1) = 1$.

**Definition 3** [36]**.** An aggregation function $T$: $[0, 1]^2 \to [0, 1]$ is said to be a t-norm if, for all $x$, $y$, $z \in [0, 1]$, the following conditions hold:

(**T1**) Commutativity: $T(x, y) = T(y, x)$;
(**T2**) Associativity: $T(x, T(y, z)) = T(T(x, y), z)$;
(**T3**) Boundary condition: $T(1, x) = T(x, 1) = x$.

**Definition 4** [10,19,21]**.** A function $O$: $[0, 1]^2 \to [0, 1]$ is an overlap function if, for all $x$, $y$, $z \in [0, 1]$, the following conditions hold:

(**O1**) $O$ is commutative;
(**O2**) $O(x, y) = 0$ if and only if $x = 0$ or $y = 0$;
(**O3**) $O(x, y) = 1$ if and only if $x = y = 1$;
(**O4**) $O$ is increasing;
(**O5**) $O$ is continuous.

**Definition 5** [3]**.** A bivariate function $C$: $[0, 1]^2 \to [0, 1]$ is said to be a copula if, for all $x$, $x'$, $y$, $y' \in [0, 1]$ with $x \le x'$ and $y \le y'$, the following conditions hold:

(**C1**) $C(x, y) + C(x', y') \ge C(x, y') + C(x', y)$;
(**C2**) $C(x, 0) = C(0, x) = 0$;
(**C3**) $C(x, 1) = C(1, x) = x$ .

Observe that overlap functions, t-norms and copulas can be extended to n-ary functions (see, e.g., [26,27,29,36]).

**Definition 6** [9]**.** 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 $\vec{r}$-increasing if, for all vectors $(x_1, \ldots, x_n) \in [0, 1]^n$ and for all $c > 0$ such that $(x_1 + cr_1, \ldots, x_n + cr_n) \in [0, 1]^n$, it holds

$$F(x_1 + cr_1, \ldots, x_n + cr_n) \ge F(x_1, \ldots, x_n). \tag{1}$$

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

**Definition 7** [22,40]**.** Let $\vec{r} = (r_1, \ldots, r_n)$ be a real $n$-dimensional vector, $\vec{r} \neq \vec{0}$. A function $PA$: $[0, 1]^n \to [0, 1]$ is said to be an n-ary pre-aggregation function if it satisfies **(A2)** and it is $\vec{r}$-increasing. We say that $PA$ is an $\vec{r}$-pre-aggregation function.

**Example 1.** In this example, we analyze the basic properties (LAE), (RNE) and (LC) for some pre-aggregation functions.

1. The function $F_{NA}$: $[0, 1]^2 \to [0, 1]$, defined by:

$$F_{NA}(x, y) = \begin{cases} x & \text{if } x \le y \\ \min\{\frac{x}{2}, y\} & \text{otherwise} \end{cases}$$

is a left 0-absorbent pre-aggregation function. In fact, it is immediate that $F_{NA}$ satisfies **(A2)**. Moreover, consider $x$, $y \in [0, 1]$ and $c > 0$ such that $y + c \in [0, 1]$. To show that $F_{NA}$ is (0, 1)-increasing, consider the following cases:

$x \le y$: In this case, it holds that $x \le y + c$. It follows that:

$$F_{NA}(x, y + c) = x = F_{NA}(x, y).$$

$x > y$: If $x > y + c$, then one has that:

$$F_{NA}(x, y + c) = \min\left\{\frac{x}{2}, y + c\right\} \ge \min\left\{\frac{x}{2}, y\right\} = F_{NA}(x, y).$$

Now suppose that $x \le y + c$. Then, it follows that:

$$F_{NA}(x, y + c) = x > \min\left\{\frac{x}{2}, y\right\} = F_{NA}(x, y).$$

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

Thus, $F_{NA}$ is a (0, 1)-pre-aggregation function. In fact, $F_{NA}$ is $\vec{r}$-increasing whenever the non-zero vector $\vec{r} = (r_1, r_2)$ satisfies $r_2 \geq r_1 \geq 0$. Hence, $F_{NA}$ is also (1, 1)-increasing. Finally, observe that $F_{NA}$ is left 0-absorbing **(LAE)**, since $F_{NA}(0, y) = 0$, for all $y \in [0, 1]$. Additionally, $F_{NA}$ satisfies **(RNE)** and **(LC)**.

2. Consider now the function $F_{NA1}$: $[0, 1]^2 \to [0, 1]$, defined by

$$F_{NA1}(x, y) = \begin{cases} \frac{x+y}{2} & \text{if } x \leq y \\ \min\{\frac{x}{2}, y\} & \text{otherwise.} \end{cases}$$

Similarly, one can show that $F_{NA1}$ is a (0, 1)-pre-aggregation function. However, it is not left 0-absorbent, since, for example $F_{NA1}(0, 0.2) = 0.1 \neq 0.2$. Moreover, $F_{NA1}$ satisfies neither **RNE** nor **LC**.

3. Consider a slight modification in the definition of the function $F_{NA1}$, obtaining the function $F_{NA2}$: $[0, 1]^2 \to [0, 1]$, defined by

$$F_{NA2}(x, y) = \begin{cases} 0 & \text{if } x = 0 \\ \frac{x+y}{2} & \text{if } 0 < x \leq y \\ \min\{\frac{x}{2}, y\} & \text{otherwise.} \end{cases}$$

Again, analogously, it is possible to show that $F_{NA2}$ is a (0, 1)-pre-aggregation function and it is immediate that $F_{NA2}$ is left 0-absorbent **(LAE)**. However, $F_{NA2}$ does not satisfy neither **RNE** nor **LC**.

4. Similarly, one can modify $F_{NA2}$ into $F_{NA3}$: $[0, 1]^2 \to [0, 1]$, defined by

$$F_{NA3}(x, y) = \begin{cases} x & \text{if } y = 1 \\ F_{NA2}(x, y) & \text{otherwise,} \end{cases}$$

and then $F_{NA3}$ satisfies all three properties **(LAE), (RNE)** and **(LC)**. However, although $F_{NA3}$ satisfies **(A2)**, it is not (0, 1) increasing, since, for example, for $x = 0.4$, $y = 0.8$ and $c = 0.2$, one has that $F_{NA3}(0.4, 0.8 + 0.2) = 0.4 <= 0.6 = \frac{0.4+0.8}{2} = F_{NA3}(0.4, 0.8)$.

Finally, it is worth mentioning that $F_{NA}$, $F_{NA1}$ and $F_{NA2}$ are all (1, 0)-increasing, but $F_{NA3}$ is not.

**Definition 8** [40, Thorem 4.1]**.** Let $\vec{r} = (r_1, \ldots, r_n)$ be a real $n$-dimensional vector, $\vec{r} \neq \vec{0}$. An $\vec{r}$-pre-aggregation function $PA$: $[0, 1]^n \to [0, 1]$ is averaging if

$$\min \leq PA \leq \max.$$

Observe that there exist pre-aggregation functions that are averaging but are not aggregation functions, for example, the mode.

**Remark 1.** Observe that all $\vec{r}$-pre-aggregation functions $PA$ that are averaging are also idempotent. However the converse does not hold. For example, consider the (0, 1)-pre-aggregation function $F_{NA}$ of Example 1, which is obviously idempotent. $F_{NA}$ is not averaging, since, for example:

$$F_{NA}(0.5, 0.4) = \min\{0.25, 0.4\} = 0.25 < \min\{0.5, 0.4\}.$$

Fuzzy integrals are well known aggregation operators. However, their use is not easy as their interpretation is not straightforward. In [54], Torra and Narukawa study the interpretation of fuzzy integrals, focusing on Sugeno ones, showing their application in fuzzy inference systems when the rules are not independent, for control problems.

The Choquet integral is a type of aggregation function which considers the relationship among the elements that are being aggregated, providing the relevance of a coalition by fuzzy measures.

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

**Definition 9** [13,47]**.** 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 $\mathfrak{m}(N) = 1$.

In this paper, we have selected the power measure according to the results in [4,11,39,40,43]. The power measure is defined as $\mathfrak{m}_{PM} : 2^N \to [0, 1]$, which is given, for all $X \subseteq N$, by

$$\mathfrak{m}_{PM}(X) = \left(\frac{|X|}{n}\right)^q, \text{ with } q > 0. \tag{2}$$

**Definition 10** [8, Definition 1.74]**.** 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:

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

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 $\vec{x}$.

## 3. Construction of pre-agregation functions using Choquet integrals and left 0-absorbent functions

In [40], we introduced the concept of a pre-aggregation function, presenting a construction method of idempotent and averaging pre-aggregation functions by means of the Choquet integral. To do it, we replace the product operation in Eq. (3) by functions $F$ that are (1, 0)-pre-aggregation functions satisfying **(LAE), (RNE)** and **(LC)** (see [40, Theorem 4.1]). The application shown in that paper considered just the case when $F$ is a t-norm, which obviously satisfies those three properties.

In this paper, we intend to propose a more general way for this construction, since we do not require $F$ to be an (1, 0)-pre-aggregation function. That is, just the conditions **(LAE)** and **(RNE)** are necessary to have idempotent pre-aggregation functions. In case we want to obtain also averaging pre-aggregation functions the functions F also have to fulfill the **(LC)** property.

In the following, we present the method for constructing a family of pre-aggregation functions defined by generalizing the discrete Choquet Integral using left 0-absorbent functions $F$: $[0, 1]^2 \to [0, 1]$, obtaining the so-called $C_F$-integrals.

**Definition 11.** Let $F$: $[0, 1]^2 \to [0, 1]$ be a bivariate function and $\mathfrak{m} : 2^N \to [0,1]$ be a fuzzy measure. The Choquet-like integral based on $F$ with respect to $\mathfrak{m}$, called $C_F$-integral, is the function $\mathfrak{C}^F_\mathfrak{m} : [0,1]^n \to [0,1]$, defined, for all $x \in [0, 1]^n$, by

$$\mathfrak{C}^F_\mathfrak{m}(\vec{x}) = \min\left\{1, \sum_{i=1}^{n} F\Big(x_{(i)} - x_{(i-1)}, \mathfrak{m}\Big(A_{(i)}\Big)\Big)\right\}, \tag{4}$$

where $(x_{(1)}, \ldots, x_{(n)})$ is an increasing permutation on the input $\vec{x}$, that is, $0 \le x_{(1)} \le \ldots \le x_{(n)}$, 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 $\vec{x}$.

**Proposition 1.** *$\mathfrak{C}^F_\mathfrak{m}$ is well defined, for any function $F$: $[0, 1]^2 \to [0, 1]$ and fuzzy measure $\mathfrak{m} : 2^N \to [0,1]$.*

**Proof.** It is immediate. □

**Remark 2.** There are some other approaches presenting Choquet-like integrals or generalizations of the Choquet integrals, mostly not restricted to discrete domains. In [45], Mesiar introduced some Choquet-like integrals defined in terms of pseudo-addition and pseudo-multiplication, presenting similar properties than those of the standard Choquet Integral. Murofushi and Sugeno [46] defined the fuzzy t-conorm integral, which is a generalization of Sugeno integral and Choquet integral based on a t-system composed by continuous t-conorms and a continuous t-norm which all are either idempotent (then the Sugeno integral is obtained), or all are Archimedean (then a transform of the Choquet integral is obtained). Differently, our Choquet-like integrals, introduced in Definition 11, are obtained in the context of the discrete Choquet integral. They are based on the standard summation + (i.e., in this item less general than the two above mentioned types of integrals) and on a rather general function $F$ (much more general than the pseudo-multiplications considered in the two above integrals).

**Remark 3.** In the literature, there exist also other kinds of integrals not defined in terms of the Choquet/Sugeno integral but related to them. For example, Wang et al. [56] introduced a nonlinear integral with respect to set functions vanishing at the empty set which need not be monotone. Observe that if a fuzzy measure $\mathfrak{m}$ is considered, then this integral coincides with the concave integral introduced by Lehrer [37] (see also [38]). This integral is just the Choquet integral whenever the considered fuzzy measure $\mathfrak{m}$ is superadditive, i.e., if $\mathfrak{m}(E_1 \cup E_2) + \mathfrak{m}(E_1 \cap E_2) \ge \mathfrak{m}(E_1) + m(E_2)$, for any sets $E_1$, $E_2 \subseteq N$. In particular, these equalities hold when $\mathfrak{m}$ is a belief measure [55]. We point out that Wang et al.'s integral coincides with our $C_F$-integral only if $F$ is the standard product and $\mathfrak{m}$ is a supermodular fuzzy measure (and then they are just the standard Choquet integral, as the authors showed in [56, Corollary 2]).

**Proposition 2.** *For any fuzzy measure $\mathfrak{m} : 2^N \to [0,1]$ and left 0-absorbent **(RNE)**-function $F$: $[0, 1]^2 \to [0, 1]$, $\mathfrak{C}^F_\mathfrak{m}$ is idempotent.*

**Proof.** Considering $\vec{x} = (x, \ldots, x) \in [0,1]^n$, one has that:

$$\begin{aligned}\mathfrak{C}^{(F)}_\mathfrak{m}(\vec{x}) &= \min\left\{1, F(x-0,1) + \sum_{i=2}^{n} F\Big(x - x, \mathfrak{m}(A_{(i)})\Big)\right\} \text{ by Eq. (4)}\\ &= \min\{1, x\} \quad \text{by } (\textbf{RNE}) \text{ and } (\textbf{LAE})\\ &= x.\end{aligned}$$

□

**Proposition 3.** *For any fuzzy measure $\mathfrak{m} : 2^N \to [0,1]$ and $F$: $[0, 1]^2 \to [0, 1]$ satisfying **(RNE)**, it holds that $\mathfrak{C}^F_\mathfrak{m} \ge \min$.*

**Proof.** Let $(x_{(1)}, \ldots, x_{(i-1)}, x_{(i)}, \ldots, x_{(n)})$ be an increasing permutation of $\vec{x} \in [0, 1]^n$. It follows that:

$$\begin{aligned}
\mathfrak{C}_{\mathfrak{m}}^{F}(\vec{x}) &= \min\left\{1, F(x_{(1)} - 0, 1) + \sum_{i=2}^{n} F(x_{(i)} - x_{(i-1)}, \mathfrak{m}(A_{(i)}))\right\} \text{ by Eq. (4)} \\
&= \min\left\{1, x_{(1)} + \sum_{i=2}^{n} F(x_{(i)} - x_{(i-1)}, \mathfrak{m}(A_{(i)}))\right\} \text{ by}(\textbf{RNE}) \\
&\geq x_{(1)} \\
&= \min \vec{x}.
\end{aligned}$$

□

**Proposition 4.** *For any fuzzy measure* $\mathfrak{m} : 2^N \to [0, 1]$ *and F:* $[0, 1]^2 \to [0, 1]$ *satisfying* ***(LC)****, it holds that* $\mathfrak{C}_{\mathfrak{m}}^{F} \leq \max$.

**Proof.** Let $(x_{(1)}, \ldots, x_{(i-1)}, x_{(i)}, \ldots, x_{(n)})$ be an increasing permutation of $\vec{x} \in [0, 1]^n$. It follows that:

$$\begin{aligned}
\mathfrak{C}_{\mathfrak{m}}^{F}(\vec{x}) &= \min\left\{1, \sum_{i=1}^{n} F(x_{(i)} - x_{(i-1)}, \mathfrak{m}(A_{(i)}))\right\} \text{ by Eq. (4)} \\
&\leq \min\left\{1, \sum_{i=1}^{n} \left(x_{(i)} - x_{(i-1)}\right)\right\} \text{ by}(\textbf{LC}) \\
&= \min\{1, x_{(n)}\} \\
&= x_{(n)} \\
&= \max \vec{x}.
\end{aligned}$$

□

**Proposition 5.** *For any fuzzy measure* $\mathfrak{m} : 2^N \to [0, 1]$ *and left 0-absorbent function F:* $[0, 1]^2 \to [0, 1]$*, if F satisfies* ***(A2, ii)****, then* $\mathfrak{C}_{\mathfrak{m}}^{F}$ *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:

$$\begin{aligned}
\mathfrak{C}_{\mathfrak{m}}^{F}(\vec{0}) &= \min\left\{1, \sum_{i=1}^{n} F(0 - 0, \mathfrak{m}(A_{(i)}))\right\} \text{ by Eq. (4)} \\
&= 0 \quad \text{by}(\textbf{LAE})
\end{aligned}$$

and

$$\begin{aligned}
\mathfrak{C}_{\mathfrak{m}}^{F}(\vec{1}) &= \min\{1, F(1 - 0, \mathfrak{m}(A_{(1)})) + \sum_{i=2}^{n} F(1 - 1, \mathfrak{m}(A_{(i)}))\} \text{ by Eq. (4)} \\
&= \min\{1, F(1, 1) + \sum_{i=2}^{n} F(0, \mathfrak{m}(A_{(i)}))\} \\
&= 1. \quad \text{by } (\textbf{A2})(\textbf{ii}) \text{ and } \textbf{LAE}
\end{aligned}$$

□

**Proposition 6.** *For any fuzzy measure* $\mathfrak{m} : 2^N \to [0, 1]$*, if the function F:* $[0, 1]^2 \to [0, 1]$ *satisfies one of the following conditions:*

(i) *F is* (1, 0)*-increasing*
(ii) *F satisfies* ***(RNE)***

*then* $\mathfrak{C}_{\mathfrak{m}}^{F}$ *is* $\vec{1}$*-increasing.*

**Proof.** Let $(x_{(1)}, \ldots, x_{(i-1)}, x_{(i)}, \ldots, x_{(n)})$ be an increasing permutation of $\vec{x} \in [0,1]^n$. Suppose that **(i)** holds and consider $c > 0$ such that $\vec{x} + c \in [0,1]^n$. Then, it follows that:

$$\begin{aligned}
\mathfrak{C}^F_{\mathfrak{m}}(x_1 + c, \ldots, x_n + c) &= \min\left\{1, F(x_{(1)} + c - 0, \mathfrak{m}(A_{(1)})) + \sum_{i=2}^{n} F(x_{(i)} + c - (x_{(i-1)} + c), \mathfrak{m}(A_{(i)}))\right\} \\
&\quad \text{by Eq. (4)} \\
&\geq \min\left\{1, F(x_{(1)}, \mathfrak{m}(A_{(1)})) + \sum_{i=2}^{n} F(x_{(i)} - x_{(i-1)}, \mathfrak{m}(A_{(i)}))\right\} \text{ by}(\mathbf{i}) \\
&= \mathfrak{C}^F_{\mathfrak{m}}(x_1, \ldots, x_n). \text{ by Eq. (4)}
\end{aligned}$$

Now consider that **(ii)** holds. It follows that:

$$\begin{aligned}
\mathfrak{C}^F_{\mathfrak{m}}(x_1 + c, \ldots, x_n + c) &= \min\left\{1, F(x_{(1)} + c - 0, 1) + \sum_{i=2}^{n} F(x_{(i)} + c - (x_{(i-1)} + c), \mathfrak{m}(A_{(i)}))\right\} \\
&\quad \text{by Eq. (4)} \\
&= \min\left\{1, x_{(1)} + c + \sum_{i=2}^{n} F(x_{(i)} - x_{(i-1)}, \mathfrak{m}(A_{(i)}))\right\} \text{ by}(\mathbf{ii}) \\
&\geq \min\left\{1, x_{(1)} + \sum_{i=2}^{n} F(x_{(i)} - x_{(i-1)}, \mathfrak{m}(A_{(i)}))\right\} \\
&= \mathfrak{C}^F_{\mathfrak{m}}(x_1, \ldots, x_n). \quad \text{by Eq. (4)}
\end{aligned}$$

□

**Theorem 1.** *For any fuzzy measure* $\mathfrak{m} : 2^N \to [0,1]$ *and left 0-absorbent* ***(RNE)****-function* $F$: $[0, 1]^2 \to [0, 1]$, $\mathfrak{C}^F_{\mathfrak{m}}$ *is a* $\vec{1}$*-pre-aggregation function.*

**Proof.** It follows from Propositions 5 and 6, observing that the property **(RNE)** implies **(A2)**(ii). □

**Corollary 1.** *For any fuzzy measure* $\mathfrak{m} : 2^N \to [0,1]$ *and left 0-absorbent* ***(RNE)****-function* $F$: $[0, 1]^2 \to [0, 1]$ *satisfying* ***(LC)****,* $\mathfrak{C}^F_{\mathfrak{m}}$ *is an idempotent averaging* $\vec{1}$*-pre-aggregation function.*

**Proof.** It follows from Propositions 3 and 4, and Theorem 1. □

**Remark 4.** Observe that, even when a left 0-absorbent function $F$: $[0, 1]^2 \to [0, 1]$ is not an averaging function, we may obtain an averaging $C_F$-integral. For example, consider the left 0-absorbent function $F_{NA}$: $[0, 1]^2 \to [0, 1]$ of Example 1. By Remark 1, we know that that $F_{NA}$ is idempotent but not averaging. However, it is immediate that $F_{NA}$ satisfies (RNE) and (LC), and, therefore, by Corollary 1, the $C_F$-integral $\mathfrak{C}^{F_{NA}}_{\mathfrak{m}}$, for a fuzzy measure $\mathfrak{m}$, is an averaging idempotent $\vec{1}$-pre-aggregation function.

**Theorem 2.** *For any fuzzy measure* $\mathfrak{m} : 2^N \to [0,1]$ *and left 0-absorbent* $(1, 0)$*-pre-aggregation function* $F$: $[0, 1]^2 \to [0, 1]$, $\mathfrak{C}^F_{\mathfrak{m}}$ *is a* $\vec{1}$*-pre-aggregation function.*

**Proof.** It follows from Propositions 5 and 6, observing that any left 0-absorbent pre-aggregation function satisfies **(LAE)**. □

In Table 1 we show a set of bivariate functions $F$: $[0, 1]^2 \to [0, 1]$ that belong to different families like (I) t-norms, (II) overlap functions, (III) copulas that are neither t-norms nor overlap functions, (IV) aggregation functions not included in (I)-(III) and (V) left-0 absorbent (0, 1)-pre-aggregation functions. For each function $F$, we show its definition and reference (if they are new this field is left empty), as well as whether or not they satisfy **(LAE), (RNE), (LC), (A2)** and (1, 0)-increasingness. Then, in the last but two column, according to properties analyzed in the previous columns, we indicate whether or not the obtained $C_F$-integral (constructed using Eq. (4)) is a $\vec{1}$-pre-aggregation function (PA). Observe that the set of conditions that $F$ should fulfill for the $C_F$-integral to be a pre-aggregation function is one of the following ones:

- Theorem 1 (**(LAE)** and **(RNE)**).
- Theorem 2 (**(LAE), (A2)**, (1, 0)-increasingness).

Finally, the last but one column shows if the obtained $C_F$-integral is averaging **(AV)** (that is, if it satisfies Propositions 3 and 4), and the last column if it is idempotent **(ID)** (that is, if it satisfies Proposition 2).

**Table 1**
Analysis of the conditions of Theorems 1 and 2, Propositions 2–4, for families of left 0-absorbent functions $F$.

| Definition | Name/Reference | (LAE) | (RNE) | LC | (A2) | (1, 0)-inc. | PA | AV | ID |
|---|---|---|---|---|---|---|---|---|---|
| **(I) T-norms [36]** | | | | | | | | | |
| $T_M(x,y)=\min\{x,y\}$ | Minimum | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| $T_P(x,y)=xy$ | Algebraic Product | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| $T_L(x,y)=\max\{0,x+y-1\}$ | Łukasiewicz | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| $T_{HP}(x,y)=\begin{cases}0 & \text{if } x=y=0\\ \frac{xy}{x+y-xy} & \text{otherwise}\end{cases}$ | Hamacher Product | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| **(II) Overlap functions [5,10,19,21]** | | | | | | | | | |
| $O_B(x,y)=\min\{x\sqrt{y},y\sqrt{x}\}$ | [10, Theorem 8] Cuadras-Augé copula [48] | ✓ | | ✓ | ✓ | ✓ | ✓ | | |
| $O_{mM}(x,y)=\min\{x,y\}\max\{x^2,y^2\}$ | [19, Example 3.1.(i)], [18, Example 4] [21, Example 3.1] | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| $O_\alpha(x,y)=xy(1+\alpha(1-x)(1-y))$, $\alpha\in[-1,0[\,\cup\,]0,1]$ | [3, Apendix A (A.2.1)],[39] Farlie-Gumbel-Morgenstern copula family* | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| $O_{Div}(x,y)=\frac{xy+\min\{x,y\}}{2}$ | [3, Apendix A (A.8.7)], [41, Table 1] | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| $GM(x,y)=\sqrt{xy}$ | Geometric Mean [27, Example 1] | ✓ | | | ✓ | ✓ | ✓ | | |
| $HM(x,y)=\begin{cases}0 & \text{if } x=0 \text{ or } y=0\\ \frac{2}{\frac{1}{x}+\frac{1}{y}} & \text{otherwise}\end{cases}$ | Harmonic Mean [27, Example 1] | ✓ | | | ✓ | ✓ | ✓ | | |
| $S(x,y)=\sin\left(\frac{\pi}{2}(xy)^{\frac{1}{4}}\right)$ | Sine [27, Example 1] | ✓ | | | ✓ | ✓ | ✓ | | |
| $O_{RS}(x,y)=\min\left\{\frac{(x+1)\sqrt{y}}{2},y\sqrt{x}\right\}$ | | ✓ | | | ✓ | ✓ | ✓ | | |
| **(III) Copulas that are neither t-norms nor overlap functions [3]** | | | | | | | | | |
| $C_F(x,y)=xy+x^2y(1-x)(1-y)$ | [36, Example 9.5 (v)], [41, Table 1] | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| $C_L(x,y)=\max\{\min\{x,\frac{y}{2}\},x+y-1\}$ | [3, Apendix A (A.5.3a)], [41, Table 1] | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| **(IV) Aggregation functions other than (I)–(III)** | | | | | | | | | |
| $F_{GL}(x,y)=\sqrt{\frac{x(y+1)}{2}}$ | | ✓ | | | ✓ | ✓ | ✓ | | |
| $F_{BPC}(x,y)=xy^2$ | [8, Example 1.80] | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| **(V) Left 0-absorbent (0, 1)-pre-aggregation functions** | | | | | | | | | |
| $F_{BD1}(x,y)=\min\{x,1-x+\min\{x,y^q\}\}$, $0<q\le 1$ | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| $F_{NA}(x,y)=\begin{cases}x & \text{if } x\le y\\ \min\{\frac{x}{2},y\} & \text{otherwise}\end{cases}$ | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| $F_{NA2}(x,y)=\begin{cases}0 & \text{if } x=0\\ \frac{x+y}{2} & \text{if } 0<x\le y\\ \min\{\frac{x}{2},y\} & \text{otherwise}\end{cases}$ | | ✓ | | | ✓ | ✓ | ✓ | | ✓ |

* When $\alpha=0$, we have that $O_\alpha=T_P$, the product t-norm, which was considered in the first part of table.

## 4. Applying $C_F$-integrals in fuzzy rule-based classification systems

In this section, we firstly present the application of $C_F$-integrals in classification problems [25], adopting it to aggregate the information given by the fuzzy rules. To do so, consider that a classification problem consists of $m$ training examples, $\mathbf{x}_p = (x_{p1}, \ldots, x_{pn}, y_p)$, with $p = 1, \ldots, m$, where $x_{pi}$, with $i = 1, \ldots, n$, is the value of the $i$th attribute and $y_p \in \mathbb{C} = \{C_1, \ldots, C_M\}$ is the label of the class of the $p$th training example, where M is the number of classes.

In this work, we use FRBCSs to tackle this kind of problems. Specifically, we have selected FARC-HD [2] to accomplish the learning of the fuzzy rules, since it is one of the most precise fuzzy classifiers nowadays. The form of the fuzzy rules used by this algorithm is:

$$\text{Rule } R_j: \text{ If } x_{p1} \text{ is } A_{j1} \text{ and } \ldots \text{ and } x_{pn} \text{ is } A_{jn} \text{ then } x_p \text{ is } C_j \text{ with } RW_j,$$

where $x_p = (x_{p1}, \ldots, x_{pn})$ is the n-dimensional vector of attribute values corresponding to an example $\mathbf{x}_p$, $R_j$ is the label of the $j$th rule, $A_{ji}$ is an antecedent fuzzy set modeling a linguistic term, $C_j$ is the label of the class of the rule $R_j$, with $C_j \in \{1, \ldots, M\}$ and $RW_j \in [0, 1]$ is the rule weight [34], which, in this case, is computed using the certainty factor.

We have used the set up suggested by the authors of FARC-HD, which is as follows: the product t-norm as the conjunction operator, five linguistic labels per variable, modeled by triangular shaped membership functions, the minimum support is set to 0.05, the threshold for the confidence is 0.8 and the maximum depth of the search tree is limited to 3.

In this paper, we propose a new FRM, where $C_F$-integrals are used to obtain the information associated with each class of the problem, that is, to aggregate the local information given by the fired rules of the system when classifying a new example, $x_p$. Specifically, the predicted class for a new example $x_p$ is computed by:

$$class = \arg \max_{k=\{1,\ldots,M\}} (\mathfrak{C}^F_{m_k}(\mu_{A_j}(x_p) * RW_j \mid Class(R_j) = k)) \text{ with, } j = 1, \ldots, L. \tag{5}$$

where $\mathfrak{C}^F_{m_k}$ is the $C_F$-integral (associated with the fuzzy measure $m_k$) considered to aggregate information given by the fired rules for the class $k$, $\mu_{A_j}$ is the matching degree of the example $x_p$ with the antecedent of the $j$th fuzzy rule, $RW_j$ is its rule weight and $L$ is the number of fuzzy rules in the system.

From Eq. (5) it can be observed that we consider a different $C_F$-integral for each class of the problem. This is due to the fact that we construct a different fuzzy measure for each class of the problem. Specifically, we use the power measure (see Eq. (2)) in which a different $q$ exponent is learnt for each class of the problem using a genetic algorithm, as we have done in our previous papers in the topic (see [4,40,41] for details of the evolutionary algorithm). Regarding the parameters of this genetic algorithm, we consider a population composed of 50 individuals, 20.000 evaluations and 30 bits for each gene in the gray codification.

In the remainder of this section we present the experimental framework used to test the quality of our new FRM. In first place we present the considered datasets (Ssection 4.1) followed by the statistical tests that are used in this paper for performing comparisons (Section 4.2).

### 4.1. Datasets

In this paper, to analyze the performance of our proposal, we consider 33 different datasets selected from the KEEL[2] dataset repository [1]. The properties of the selected datasets are summarized in Table 2, showing for each dataset the identification of this dataset (ID), followed by the name of the dataset (Dataset), the number of examples (#Ex.), the number of attributes (#Atts.) and the number of classes (#Class).

Some datasets, namely: *magic, page-blocks, penbased, ring, satimage* and *twonorm*, were stratified sampled at 10% in order to reduce their size for training. Some examples containing missing information were removed, e.g., in the *wisconsin* dataset.

We have applied a 5-fold cross-validation technique, that is, split the dataset into five partitions randomly. Each partition has 20% of the examples. We use four partitions for training, and the other is used for testing. This process is repeated five times, using a different partition for testing each time. In each iteration we measure the quality of the classifier using the accuracy rate, which is defined as the number of correctly classified examples divided by the total number of examples for each partition. We then compute the average result of the five testing partitions, which is the output of the algorithm.

### 4.2. Statistical tests for performance comparisons

For the statistical analysis of the results, we use hypothesis validation techniques [28,53], namely, non-parametric tests, since the initial conditions that guarantee the reliability of the parametric tests cannot be guaranteed [15].

We apply the aligned Friedman rank test [31] to detect statistical differences among a group of results and to verify the quality of a method in comparison to its partners. The algorithm achieving the lowest average ranking is the best one.

Additionally, to analyze if the best ranking method rejects the equality hypothesis with respect to its partners we use the post-hoc Holm's test [32]. This method allows us to see whenever a hypothesis of comparison could be rejected at a specified level of significance $\alpha$. We compute the adjusted $p$-value (APV) to take into account that multiple tests are

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

**Table 2**
Summary of the properties of the datasets considered in this study.

| Id. | Dataset | #Ex. | #Atts. | #Class |
|---|---|---|---|---|
| App | Appendiciticis | 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 | 1902 | 10 | 2 |
| New | Newthyroid | 215 | 5 | 3 |
| Pag | Pageblocks | 5472 | 10 | 5 |
| Pen | Penbased | 10,992 | 16 | 10 |
| Pho | Phoneme | 5404 | 5 | 2 |
| Pim | Pima | 768 | 8 | 2 |
| Rin | Ring | 740 | 20 | 2 |
| Sah | Saheart | 462 | 9 | 2 |
| Sat | Satimage | 6435 | 36 | 7 |
| Seg | Segment | 2310 | 19 | 7 |
| Shu | Shuttle | 58,000 | 9 | 7 |
| Son | Sonar | 208 | 60 | 2 |
| Spe | Spectfheart | 267 | 44 | 2 |
| Tit | Titanic | 2201 | 3 | 2 |
| Two | Twonorm | 740 | 20 | 2 |
| Veh | Vehicle | 846 | 18 | 4 |
| Win | Wine | 178 | 13 | 3 |
| Wis | Wisconsin | 683 | 11 | 2 |
| Yea | Yeast | 1484 | 8 | 10 |

performed. Then, we can directly compare the APV with the level of significance $\alpha$, and, thus, we are able to reject the null hypothesis.

## 5. Experimental study and results

In this section, we present the results achieved in testing when using the FRM generalized by $C_F$-integrals. To do so, the results of our proposal are analyzed considering three main steps:

1. Firstly, we present the results and analyze the performance of each averaging $C_F$-integral. After that, we compare them among themselves in order to discover which generalization is the one that best represent the family of averaging $C_F$-integrals.
2. The second part of the study is related to $C_F$-integrals that are not averaging. In order to find the best representative method of this family, we firstly analyze the achieved results and after that, compare them among themselves.
3. Once we have found the two best $C_F$-integrals obtained in the previous steps, in order to test the quality of our approach, we perform the following comparisons:
   (a) The best averaging $C_F$-integral versus classical averaging functions (FRM of the WR) and our previous averaging pre-aggregation functions.
   (b) The best non-averaging $C_F$-integral versus the classical non averaging functions, like the FRM of AC or the usage of the probabilistic sum.
   (c) Finally, we test whether the application of the best non averaging function enhances the results of the averaging operators or not.

### 5.1. Analysis of the performance of averaging $C_F$-integrals

This subsection is aimed at analyzing the performance of the averaging $C_F$-integrals considered in this study (see Table 1) in the FRM. The results achieved in test by the 9 averaging functions are available in Table 3, by columns. In each row we introduce the results of each dataset, highlighting the best global result in **boldface**. We also include in this table the number of datasets where each function achieves the best (#Wins) and the worst (#Loses) result, respectively.

**Table 3**
Accuracy achieved in test by different averaging $C_F$-integrals.

| Dataset | $O_\alpha$ | $O_B$ | $O_{mM}$ | $O_{Div}$ | CF | CL | $F_{BPC}$ | $F_{BD1}$ | $F_{NA}$ |
|---|---|---|---|---|---|---|---|---|---|
| App | 83.03 | 83.03 | 84.94 | 83.94 | 82.99 | **85.84** | 83.07 | 83.94 | 82.99 |
| Bal | 80.32 | 82.08 | 82.56 | 81.60 | 82.24 | 84.00 | 80.32 | **84.80** | 82.56 |
| Ban | 86.09 | 86.81 | 86.85 | 85.79 | 85.70 | 85.04 | **87.02** | 83.09 | 86.09 |
| Bnd | 68.26 | **71.83** | 67.70 | 69.68 | 69.13 | 69.66 | 66.62 | 71.09 | 69.40 |
| Bup | 63.77 | 65.51 | 66.09 | 65.51 | 65.22 | 62.32 | 66.67 | 64.06 | **67.83** |
| Cle | 55.23 | 56.24 | 54.55 | 55.90 | 52.18 | 55.54 | 54.88 | 57.92 | **57.92** |
| Con | 52.00 | 52.89 | 52.75 | **53.83** | 52.54 | 50.78 | 52.95 | 52.41 | 52.27 |
| Eco | 76.49 | 76.20 | 77.08 | 75.61 | 77.09 | 78.87 | 77.69 | **79.17** | 78.88 |
| Gla | 62.14 | **66.82** | 60.75 | 63.10 | 63.58 | 63.09 | 63.54 | 65.44 | 64.51 |
| Hab | 72.85 | 72.86 | 72.21 | 72.52 | 69.92 | 72.86 | 73.17 | 72.21 | **73.51** |
| Hay | 78.75 | 78.72 | 78.01 | **80.26** | 78.75 | 79.46 | 78.01 | 77.95 | 78.72 |
| Ion | 90.04 | 88.32 | 87.47 | 89.47 | 89.46 | 89.75 | 89.47 | 89.46 | **90.60** |
| Iri | 92.00 | 94.00 | 94.00 | 93.33 | **94.67** | 93.33 | 92.00 | 93.33 | 93.33 |
| Led | 68.00 | 68.40 | 67.60 | 68.80 | **69.00** | 68.00 | 68.80 | 68.00 | 68.60 |
| Mag | **80.18** | 79.86 | 79.97 | 79.91 | 78.81 | 79.76 | 78.86 | 79.55 | 80.02 |
| New | **95.35** | 94.88 | 94.42 | **95.35** | 94.88 | 94.42 | 94.88 | 93.49 | 93.49 |
| Pag | 93.43 | 94.52 | 93.98 | 93.97 | **94.89** | 93.61 | 94.34 | 94.34 | 93.97 |
| Pen | 90.09 | 91.09 | 89.45 | 90.82 | 90.55 | 90.27 | 90.09 | **92.55** | 91.45 |
| Pho | 83.05 | 82.92 | 83.29 | 82.81 | 83.23 | **83.88** | 82.70 | 81.96 | 82.86 |
| Pim | 75.13 | 75.38 | **76.17** | 74.48 | 75.78 | 75.52 | 73.82 | 73.56 | 75.13 |
| Rin | 89.19 | 89.32 | 90.00 | 89.86 | 89.73 | 89.46 | 88.38 | 88.78 | **90.27** |
| Sah | **71.85** | 69.48 | 70.78 | 68.18 | 69.48 | 68.39 | 71.21 | 69.70 | 68.61 |
| Sat | 79.16 | 78.54 | 79.00 | 78.23 | **80.72** | 79.16 | 78.38 | 78.70 | 78.54 |
| Seg | 92.73 | 92.51 | **93.33** | 92.77 | 92.94 | 93.20 | 92.42 | 93.16 | 92.55 |
| Shu | 97.01 | 97.29 | 97.01 | **97.84** | 97.10 | 97.01 | 97.10 | 97.15 | 96.78 |
| Son | **80.78** | 75.49 | 77.93 | 77.42 | 77.89 | 79.83 | 74.05 | 80.29 | 78.85 |
| Spe | 77.48 | 77.88 | 76.75 | 76.39 | 77.51 | 76.77 | **79.76** | 74.92 | 78.26 |
| Tit | 78.87 | 78.87 | 78.87 | 78.87 | 78.87 | 78.87 | 78.87 | **79.06** | 78.87 |
| Two | 84.73 | 83.78 | 84.46 | 85.14 | **85.68** | 85.41 | 84.19 | 85.14 | 83.92 |
| Veh | 68.21 | 67.73 | 66.55 | 67.02 | **70.33** | 67.38 | 68.20 | 69.26 | 67.97 |
| Win | 95.48 | 94.97 | **97.21** | 93.24 | 94.38 | 92.13 | 96.62 | 96.62 | 96.03 |
| Wis | 96.49 | 96.63 | 96.34 | 96.34 | 96.05 | 96.19 | 96.34 | **96.64** | 96.34 |
| Yea | 56.33 | 57.35 | 57.08 | **57.48** | 57.34 | 56.87 | **57.48** | 55.12 | 56.40 |
| Mean | 79.23 | 79.46 | 79.25 | 79.26 | 79.35 | 79.29 | 79.15 | 79.48 | **79.62** |
| #Wins | 4 | 2 | 3 | 5 | 6 | 2 | 3 | 5 | 5 |
| #Loses | 4 | 2 | 6 | 4 | 5 | 4 | 7 | 6 | 4 |

**Table 4**
Statistical analysis of the methods based on averaging $C_F$-integrals.

| Algorithm | Ranking | APV |
|---|---|---|
| $\mathbf{F}_{NA}$ | **129.31** | – |
| $O_B$ | 138.93 | 1.0 |
| $F_{BD1}$ | 143.21 | 1.0 |
| CF | 145.60 | 1.0 |
| CL | 149.71 | 1.0 |
| $O_{Div}$ | 152.74 | 1.0 |
| $O_{mM}$ | 154.01 | 1.0 |
| $F_{BPC}$ | 162.77 | 0.79 |
| $O_\alpha$ | 164.68 | 0.75 |

From the results shown in Table 3, it is possible to notice that all the averaging $C_F$-integrals, except $F_{NA}$, present a mean performance between 79.15 and 79.48. We have to highlight the function $F_{NA}$ since it achieves the best global mean and also the best accuracy rate in 5 out of the 33 datasets considered in this study. Moreover, observe that the functions $O_\alpha$, $O_{Div}$, *CF*, $F_{BD1}$ and $F_{NA}$ achieves similar results in terms of number of datasets with the best and worst performance respectively, while the remainder functions achieves worse results.

In order to select objectively the best function among this group, we have carried out a statistical study according to the recommendations made in the specialized literature [16,28,53].

Specifically, we have performed the aligned Friedman rank test to compare the 9 approaches, whose obtained rankings are presented in the second column of Table 4. In this table we sort the values from the lowest to highest obtained ranking, where the best one is highlighted in **boldface**. Then, we apply the Holm's post-hoc test, to check whether the control approach (the one associated with the best ranking) is statistically better than the remainder approaches, showing the obtained APV in the last column of this table.

**Table 5**
Accuracy achieved in test by different non-averaging $C_F$-integrals.

| Dataset | GM | HM | Sin | $O_{RS}$ | $F_{GL}$ | $F_{NA2}$ |
|---|---|---|---|---|---|---|
| App | 82.08 | 83.98 | 85.80 | 83.98 | 82.08 | **85.84** |
| Bal | 88.48 | 86.40 | **89.44** | 88.00 | 89.12 | 88.64 |
| Ban | 85.28 | **86.19** | 82.79 | 85.58 | 83.42 | 84.60 |
| Bnd | 71.30 | 70.51 | **72.69** | 69.19 | 71.01 | 70.48 |
| Bup | 61.16 | **66.96** | 63.48 | 66.09 | 62.03 | 64.64 |
| Cle | 57.23 | 57.24 | **57.55** | 55.88 | 57.25 | 56.55 |
| Con | 53.77 | 52.21 | **54.31** | 53.84 | 54.24 | 53.16 |
| Eco | 81.26 | 79.47 | **82.45** | 80.07 | 81.55 | 80.08 |
| Gla | 65.44 | **69.17** | 66.83 | 65.89 | 66.33 | 66.83 |
| Hab | 71.54 | 71.88 | 71.87 | **72.87** | 70.24 | 71.87 |
| Hay | 79.49 | 79.43 | 77.98 | **81.77** | 78.69 | 79.43 |
| Ion | **90.89** | 88.91 | 87.46 | 88.32 | 90.04 | 89.75 |
| Iri | **94.67** | 94.00 | 94.00 | 94.00 | 94.00 | 94.00 |
| Led | 68.00 | 68.40 | 69.60 | 69.20 | **68.80** | **69.80** |
| Mag | 80.02 | 80.23 | 79.34 | **80.23** | 79.70 | 79.70 |
| New | **97.67** | 95.81 | 95.35 | 95.81 | **97.67** | 96.28 |
| Pag | 94.34 | 93.97 | 94.34 | **94.52** | 94.34 | 94.15 |
| Pen | 92.18 | 92.09 | 91.45 | 92.00 | 92.73 | **92.91** |
| Pho | 82.07 | **83.73** | 80.96 | 82.72 | 81.27 | 81.44 |
| Pim | 74.87 | 74.87 | 75.13 | 75.00 | **76.82** | 74.61 |
| Rin | 90.95 | 90.00 | 88.51 | 90.27 | **91.35** | 89.86 |
| Sah | 69.04 | 68.84 | 71.20 | **71.86** | 70.33 | 70.12 |
| Sat | 79.01 | 78.69 | 77.45 | **80.72** | 78.53 | 80.41 |
| Seg | **93.46** | 92.73 | 92.47 | 92.77 | 93.07 | 92.42 |
| Shu | 97.06 | 96.92 | 96.69 | 96.37 | 96.69 | **97.15** |
| Son | 82.73 | 81.28 | 81.74 | 83.19 | **83.69** | 83.21 |
| Spe | 77.51 | 79.76 | 78.65 | 78.27 | 79.76 | **79.77** |
| Tit | **78.87** | **78.87** | **78.87** | **78.87** | **78.87** | **78.87** |
| Two | 89.19 | 86.89 | 91.49 | 89.05 | 90.00 | **92.57** |
| Veh | 67.97 | 68.79 | 64.77 | 67.38 | **69.03** | 68.08 |
| Win | 96.08 | 96.05 | **97.17** | 97.16 | 95.49 | 96.08 |
| Wis | 96.93 | **97.07** | 96.34 | 97.06 | 95.76 | 96.78 |
| Yea | 56.94 | **58.15** | 56.47 | 57.28 | 57.68 | 57.08 |
| Mean | 80.23 | 80.29 | 80.14 | 80.46 | 80.35 | **80.52** |
| #Wins | 5 | 8 | 7 | 7 | 7 | 7 |
| #Loses | 4 | 7 | 11 | 3 | 5 | 2 |

From these results, it is noticeable that there are no statistical differences among the averaging functions. However, for the sake of selecting a representative for this family, we choose the function $F_{NA}$ as it obtains the best global mean and it is selected as the control method in the Holm's test.

### *5.2. Studying the quality of non-averaging $C_F$-integrals*

In this subsection, we present the achieved results in testing when one considers $C_F$-integrals that do not have averaging characteristics. We present the results of the 6 functions of this type in Table 5, by columns. In each row we introduce the results of each dataset where the best result is highlighted in **boldface**. Like in Table 3, #Wins and #Loses represent the number of datasets where the function obtains the best and worst result, respectively.

From the results presented in Table 5, we can directly conclude that the non-averaging functions have a superior mean in relation to the averaging functions (Table 3), since the smallest obtained mean (80.14 by Sin) is superior than the best averaging $C_F$-integral (79.62 by $F_{NA}$). Additionally, we have to highlight the leap in performance provided by the usage of $F_{NA2}$ and $O_{RS}$. The first function has the best accuracy in 7 datasets and the worst accuracy in only 2 dataset. The function $O_{RS}$ also achieves a good mean, with best accuracy in 7 datasets and the worst one in 3 datasets. Furthermore, we should stress that although the number of dataset in which the remainder functions provide the best results is similar the number of datasets where they provide the worst result is larger, which implies a decrease on the overall performance as shown in Table 5.

According to the obtained results, it is necessary to conduct a statistical analysis to select the best function among this group. In order to do it, we have performed the same statistical study as in the previous section. The results of the aligned Friedman and Holm's tests are presented in Table 6. As expected, according to Table 5, all methods present a similar behavior, therefore, we select $F_{NA2}$ as representative of this family since it is considered as control variable and it also achieves the best global mean.

**Table 6**
Average rankings of the non-averaging $C_F$-integrals (aligned Friedman).

| Algorithm | Ranking | APV |
|---|---|---|
| $\mathbf{F_{NA2}}$ | **91.78** | - |
| $F_{GL}$ | 95.12 | 1.0 |
| $O_{RS}$ | 95.39 | 1.0 |
| GM | 99.22 | 1.0 |
| HM | 105.13 | 1.0 |
| Sin | 110.33 | 0.94 |

**Table 7**
Results achieved in test by the averaging FRMs.

| Dataset | $F_{NA}$ | Choquet | $Ham_{PA}$ | $CP_{Min}$ | WR |
|---|---|---|---|---|---|
| App | 82.99 | 80.13 | 82.99 | **85.84** | 83.03 |
| Bal | 82.56 | 82.40 | **82.72** | 81.60 | 81.92 |
| Ban | 86.09 | **86.32** | 85.96 | 84.30 | 83.94 |
| Bnd | 69.40 | 68.56 | **72.13** | 71.06 | 69.40 |
| Bup | **67.83** | 66.96 | 65.80 | 61.45 | 62.03 |
| Cle | **57.92** | 55.58 | 55.58 | 54.88 | 56.91 |
| Con | 52.27 | 51.26 | **53.09** | 52.61 | 52.07 |
| Eco | 78.88 | 76.51 | **80.07** | 77.09 | 75.62 |
| Gla | 64.51 | 64.02 | 63.10 | **69.17** | 64.99 |
| Hab | 73.51 | 72.52 | 72.21 | **74.17** | 70.89 |
| Hay | 78.72 | 79.49 | 79.49 | **81.74** | 78.69 |
| Ion | **90.60** | 90.04 | 89.18 | 88.89 | 90.03 |
| Iri | 93.33 | 91.33 | 93.33 | 92.67 | **94.00** |
| Led | 68.60 | 68.20 | 68.60 | 68.40 | **69.40** |
| Mag | **80.02** | 78.86 | 79.76 | 79.81 | 78.60 |
| New | 93.49 | 94.88 | **95.35** | 93.95 | 94.88 |
| Pag | 93.97 | 94.16 | **94.34** | 93.97 | 94.16 |
| Pen | **91.45** | 90.55 | 90.82 | 91.27 | **91.45** |
| Pho | 82.86 | 82.98 | **83.83** | 82.94 | 82.29 |
| Pim | 75.13 | 74.60 | 73.44 | **75.78** | 74.60 |
| Rin | 90.27 | **90.95** | 88.78 | 87.97 | 90.00 |
| Sah | 68.61 | 69.69 | 70.77 | **70.78** | 68.61 |
| Sat | 78.54 | 79.47 | **80.40** | 79.01 | 79.63 |
| Seg | 92.55 | **93.46** | 93.33 | 92.25 | 93.03 |
| Shu | 96.78 | 97.61 | 97.20 | **98.16** | 96.00 |
| Son | 78.85 | 77.43 | **79.34** | 76.95 | 77.42 |
| Spe | 78.26 | 77.88 | 76.02 | **78.99** | 77.90 |
| Tit | **78.87** | **78.87** | **78.87** | **78.87** | **78.87** |
| Two | 83.92 | 84.46 | 85.27 | 85.14 | **86.49** |
| Veh | 67.97 | 68.44 | 68.20 | **69.86** | 66.67 |
| Win | 96.03 | 93.79 | **96.63** | 93.83 | 96.60 |
| Wis | 96.34 | **97.22** | 96.78 | 95.90 | 96.34 |
| Yea | 56.40 | 55.73 | 56.53 | **57.01** | 55.32 |
| Mean | 79.62 | 79.22 | **79.69** | 79.58 | 79.15 |
| #Wins | 6 | 5 | 11 | 11 | 5 |
| #Loses | 4 | 7 | 3 | 8 | 10 |

### *5.3. Comparisons of the best $C_F$-integrals against classical FRMs*

Once we have selected the functions that represent the family of $C_F$-integrals with averaging or non-averaging characteristics ($F_{NA}$ and $F_{NA2}$), we compare them against classical averaging (Section 5.3.1) and non-averaging functions (Section 5.3.2), respectively.

#### *5.3.1. Analyzing the behavior of the representative averaging $C_F$-integral*

In first place, we compare the best averaging function against FRMs where averaging aggregations are applied. Namely, the FRM of the Winning Rule (WR) [8], the standard Choquet integral (Choquet) [4], the best pre-aggregation function presented in [40] (which is named $Ham_{PA}$ since it is based on the Hamacher t-norm) and the best Choquet-Like Copula-based [41] (which is named $CP_{Min}$ as it is based on the Minimum t-norm). We have to point out that the pre-aggregation function based on the Hamacher t-norm ($Ham_{PA}$) is also a $C_F$-integral (where $F$ is the Hamacher t-norm).

The results achieved in test by this averaging FRMs are available in Table 7, using the same structure as the tables presented before.

**Table 8**
Statistical analysis of the FRMs based on averaging operators.

| Algorithm | Ranking | APV |
|---|---|---|
| **Ham$_{PA}$** | **68.96** | |
| $F_{NA}$ | 76.25 | 0.62 |
| CP$_{Min}$ | 80.87 | 0.62 |
| Choquet | 92.98 | 0.12 |
| WR | 95.90 | 0.08 |

**Table 9**
Results achieved in test by classical non-averaging operators.

| Dataset | $F_{NA2}$ | AC | ProbSum |
|---|---|---|---|
| App | **85.84** | 83.03 | 85.84 |
| Bal | **88.64** | 85.92 | 87.20 |
| Ban | 84.60 | **85.30** | 84.85 |
| Bnd | **70.48** | 68.28 | 68.82 |
| Bup | 64.64 | **67.25** | 61.74 |
| Cle | 56.55 | 56.21 | **59.25** |
| Con | **53.16** | 53.16 | 52.21 |
| Eco | 80.08 | **82.15** | 80.95 |
| Gla | **66.83** | 65.44 | 64.04 |
| Hab | 71.87 | **73.18** | 69.26 |
| Hay | **79.43** | 77.95 | 77.95 |
| Ion | **89.75** | 88.90 | 88.32 |
| Iri | 94.00 | 94.00 | **95.33** |
| Led | **69.80** | 69.60 | 69.20 |
| Mag | 79.70 | **80.76** | 80.39 |
| New | **96.28** | 94.88 | 94.42 |
| Pag | 94.15 | **95.07** | 94.52 |
| Pen | 92.91 | 92.55 | **93.27** |
| Pho | 81.44 | 81.70 | **82.51** |
| Pim | 74.61 | 74.74 | **75.91** |
| Rin | 89.86 | **90.95** | 90.00 |
| Sah | **70.12** | 68.39 | 69.69 |
| Sat | **80.41** | 79.47 | 80.40 |
| Seg | 92.42 | **93.12** | 92.94 |
| Shu | **97.15** | 95.59 | 94.85 |
| Son | **83.21** | 78.36 | 82.24 |
| Spe | **79.77** | 77.88 | 77.90 |
| Tit | **78.87** | **78.87** | **78.87** |
| Two | **92.57** | 90.95 | 90.00 |
| Veh | 68.08 | **68.56** | 68.09 |
| Win | **96.08** | 96.03 | 94.92 |
| Wis | 96.78 | 96.63 | **97.22** |
| Yea | 57.08 | 58.96 | **59.03** |
| Mean | **80.52** | 80.12 | 80.07 |
| #Wins | 17 | 10 | 8 |
| #Loses | 11 | 12 | 11 |

From these results, it is possible to observe that WR and Choquet have a low mean while Ham$_{PA}$ is the one obtaining the best global mean, followed by our new averaging $C_F$-integral, $F_{NA}$ and $CP_{Min}$. Moreover, observe that Ham$_{PA}$ is also the function that has the biggest number of good results (along with CP$_{Min}$) and the lowest number of cases having bad results.

We have conducted the same statistical study than in the previous sections, where the achieved results are presented in Table 8. These results are sorted according to the ranking and highlighting in **boldface** the control ranking. Whenever there is an statistical difference in favor to the control method the APV is underlined.

Observing these results, we can see that the pre-aggregation function based on the Hamacher t-norm is considered as control variable, and it also presents differences against the FRM of the WR and a positive trend versus the standard Choquet integral. On the other hand, it is not possible to affirm that there are differences against the remainder functions. Therefore, we consider Ham$_{PA}$ as the best averaging $C_F$-integral since it achieves the best mean and the largest number of datasets having the best accuracy.

#### *5.3.2. Analyzing the behavior of the best non-averaging $C_F$-integral*

Next, we study the behavior of the non-averaging operators. Specifically, we compare our representative for the family of non-averaging $C_F$-integral, $F_{NA2}$, against the classical FRMs of the additive combination (AC) and the probabilistic sum (PS). The results are available in Table 9, having the same structure of our results presented before.

**Table 10**
Statistical analysis of the methods based on non-averaging operators.

| Algorithm | Ranking | APV |
|---|---|---|
| $\mathbf{F}_{NA2}$ | **41.80** | |
| AC | 53.65 | 0.14 |
| PS | 54.54 | 0.14 |

**Table 11**
Statistical analysis of the best non-averaging $C_F$-integral against the averaging operators.

| Algorithm | Ranking | APV |
|---|---|---|
| $\mathbf{F}_{NA2}$ | **64.93** | |
| $\text{Ham}_{PA}$ | 90.77 | <u>0.06</u> |
| $F_{NA}$ | 99.40 | <u>0.02</u> |
| $\text{CP}_{Min}$ | 102.95 | <u>0.02</u> |
| Cho | 117.43 | <u>7.91E-4</u> |
| WR | 121.48 | <u>3.05E-4</u> |

We can see in the obtained results that the $C_F$-integral based on the function $F_{NA2}$ presents the highest global mean. If we look closer, this function achieves the best accuracy in almost half of the datasets considered in this study. On the other hand, the classical aggregation functions applied in the FRM provide the best result in a lower number of datasets (10 and 8, respectively).

Again, to statistically compare these methods among themselves, we perform the Friedman rank test and the Holm's post hoc test. The obtained results are available in Table 10. These results show that our new function $F_{NA2}$ has the best rank and, consequently, it is considered as control variable as it was expected according to the previous results. When we observe the obtained APVs we can see that their APVs are low, which shows a trend pointing out that our new non-averaging function is very competitive versus these two classical aggregation functions. Therefore, the quality of our proposal is proved since it is enhancing the results of the classical FRM of the AC and PS.

To finish our study, for the sake of certifying the quality of our new function ($F_{NA2}$), we also compare it versus the five averaging FRMs studied in the previous Section 5.3.1. To accomplish this comparison, we have performed again the Friedman rank test and Holm's post hoc test among these approaches. The results of the statistical test and the obtained APVs are shown in Table 11, where the ranking related to the function considered as control method is highlighted in **boldface**. Furthermore, the APV is underlined when there are statistical differences favorably to the control approach versus the opponent method.

The obtained statistical results clearly show the superiority of the $C_F$-integral based on the function $F_{NA2}$, since it achieves statistical differences versus all the remainder methods. All in all, the non-averaging $C_F$-integral constructed using the function $F_{NA2}$, has proven to be the best choice among all developed functions, since it offers the best performance, it statistically outperforms classical averaging functions applied in the FRM and it is competitive with respect classical non averaging FRMs.

## 6. Conclusion

In this paper we have proposed a generalization of the Choquet integral by replacing its product operator by a function $F$ with some weak properties. As a result, we have defined the $C_F$-integrals, a new family of pre-aggregation functions with some particular characteristics, which allows us to enlarge the scope of the methodology that we proposed in [40]. The main advantages of this approach in relation to our previous work concerning generalizations of the Choquet integral are:

- The function $F$ used for the generalization may satisfy a less number of properties, and we still have a pre-aggregation function.
- The resulting pre-aggregation function does not need to be neither an averaging nor idempotent function.

We have applied these averaging and non-averaging $C_F$-integrals in FRBCSs to tackle classification problems. Precisely, in this work we performed a study considering 33 different public datasets, and the conclusions we draw are the following ones:

1. The considered averaging $C_F$-integrals present a similar performance than that of our previous generalizations.
2. The best averaging $C_F$-integral is $\text{Ham}_{PA}$, which was previously introduced in another paper. However, it is also a $C_F$-integral (based on the Hamacher t-norm).
3. The non-averaging $C_F$-integrals, as expected, offer a performance superior than the averaging ones.
4. The best $C_F$-integral, $F_{NA2}$, provides results that are statistically superior than all classical FRMs, and also, very competitive with the classical non-averaging FRMs like AC or PS.

Consequently, we have created a new family of pre-aggregation functions, which provides accurate results when they have non-averaging features.

Future work is concerned with two lines of research. In one hand, we will search for a generalization of our CC-integrals [41], by means of two arbitrary functions $F_1$ and $F_2$, to put in the place of each copula, satisfying a minimal set of properties that guarantee that the generalized CC-integral is, at least, a pre-aggregation function. On the other hand, we will study our generalizations in an interval-valued context, following the approach in [6,7,17], as in [49–51].


## Acknowledgments

This work is supported by Brazilian National Counsel of Technological and Scientific Development CNPq (Proc. 233950/2014-1, 306970/2013-9, 307781/2016-0), by grant APVV-14-0013, by the Spanish Ministry of Science and Technology (under project TIN2016-77356-P (AEI/FEDER, UE)), and by Caixa and Fundación Caja Navarra of Spain.


## References


[1] 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 Comput. 13 (3) (2009) 307–318.
[2] 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 Trans. Fuzzy Syst. 19 (5) (2011) 857–872.
[3] C. Alsina, M.J. Frank, B. Schweizer, Associative Functions: Triangular Norms and Copulas, World Scientific Publishing Company, Singapore, 2006.
[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] B.C. Bedregal, G.P. Dimuro, H. Bustince, E. Barrenechea, New results on overlap and grouping functions, Inf. Sci. 249 (2013) 148–170.
[6] B.C. Bedregal, G.P. Dimuro, R.H.S. Reiser, An approach to interval-valued R-implications and automorphisms, in: J.P. Carvalho, D. Dubois, U. Kaymak, J.M.d. C. Sousa (Eds.), Proceedings of the Joint 2009 International Fuzzy Systems Association World Congress and 2009 European Society of Fuzzy Logic and Technology Conference, IFSA/EUSFLAT, 2009.
[7] B.C. Bedregal, G.P. Dimuro, R.H.N. Santiago, R.H.S. Reiser, On interval fuzzy S-implications, Inf. Sci. 180 (8) (2010) 1373–1389.
[8] G. Beliakov, A. Pradera, T. Calvo, Aggregation Functions: A Guide for Practitioners, Springer, Berlin, 2007.
[9] H. Bustince, J. Fernandez, A. Kolesárová, R. Mesiar, Directional monotonicity of fusion functions, Eur. J. Oper. Res. 244 (1) (2015) 300–308.
[10] H. Bustince, J. Fernandez, R. Mesiar, J. Montero, R. Orduna, Overlap functions, nonlinear analysis: theory, Methods Appl. 72 (3–4) (2010) 1488–1499.
[11] H. Bustince, J.A. Sanz, G. Lucca, G.P. Dimuro, B. Bedregal, R. Mesiar, A. Kolesárová, G. Ochoa, Pre-aggregation functions: definition, properties and construction methods, in: Fuzzy Systems (FUZZ-IEEE), 2016 IEEE International Conference on, IEEE, 2016.
[12] H. Bustince, J.A. Sanz, G. Lucca, G.P. Dimuro, B. Bedregal, R. Mesiar, A. Kolesárová, G. Ochoa, Pre-aggregation functions: definition, properties and construction methods, in: 2016 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), IEEE, Los Alamitos, 2016.
[13] G. Choquet, Theory of capacities, Annales de l'Institut Fourier 5 (1953-1954) 131–295.
[14] O. Cordón, M.J. del Jesus, F. Herrera, A proposal on reasoning methods in fuzzy rule-based classification systems, Int. J. Approx. Reason. 20 (1) (1999) 21–45.
[15] J. Demšar, Statistical comparisons of classifiers over multiple data sets, J. Mach. Learn. Res. 7 (2006) 1–30.
[16] 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 Evol. Comput. 1 (1) (2011) 3–18.
[17] G.P. Dimuro, On interval fuzzy numbers, 2011 Workshop-School on Theoretical Computer Science, WEIT 2011, IEEE, Los Alamitos, 2011.
[18] G.P. Dimuro, B. Bedregal, J. Fernandez, Additive generators of overlap functions, in: H. Bustince, 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.
[19] G.P. Dimuro, B. Bedregal, Archimedean overlap functions: the ordinal sum and the cancellation, idempotency and limiting properties, Fuzzy Sets Syst. 252 (2014) 39–54.
[20] G.P. Dimuro, B. Bedregal, On residual implications derived from overlap functions, Inf. Sci. 312 (2015) 78–88.
[21] G.P. Dimuro, B. Bedregal, H. Bustince, M.J. Asiáin, R. Mesiar, On additive generators of overlap functions, Fuzzy Sets Syst. 287 (2016) 76–96. Theme: Aggregation Operations
[22] G.P. Dimuro, B. Bedregal, H. Bustince, J. Fernandez, G. Lucca, R. Mesiar, New results on pre-aggregation functions, in: Uncertainty Modelling in Knowledge Engineering and Decision Making, Proceedings of the 12th International FLINS Conference (FLINS 2016), vol. 10 of World Scientific Proceedings Series on Computer Engineering and Information Science, World Scientific, Singapura, 2016, pp. 213–219.
[23] G.P. Dimuro, B. Bedregal, H. Bustince, A. Jurio, M. Baczyński, K. Miś, *QL*-operations and *QL*-implication functions constructed from tuples $(O, G, N)$ and the generation of fuzzy subsethood and entropy measures, Int. J. Approx. Reason. 82 (2017) 170–192, doi:10.1016/j.ijar.2016.12.013. ISSN 0888-613X.
[24] G.P. Dimuro, B. Bedregal, R.H.N. Santiago, On $(G, N)$-implications derived from grouping functions, Inf. Sci. 279 (2014) 1–17.
[25] R.O. Duda, P.E. Hart, D.G. Stork, Pattern Classification, 2nd, Wiley-Interscience, 2000.
[26] M. Elkano, M. Galar, J. Sanz, H. Bustince, Fuzzy rule-based classification systems for multi-class problems using binary decomposition strategies: on the influence of n-dimensional overlap functions in the fuzzy reasoning method, Inf. Sci. 332 (2016) 94–114.
[27] 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 Trans. Fuzzy Syst. 23 (5) (2015) 1562–1580.
[28] 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 Comput. 13 (10) (2009) 959–977.
[29] D. Gómez, J.T. Rodríguez, J. Montero, H. Bustince, E. Barrenechea, N-dimensional overlap functions, Fuzzy Sets Syst. 287 (2016) 57–75. Theme: Aggregation Operations
[30] M. Grabisch, J. Marichal, R. Mesiar, E. Pap, Aggregation Functions, Cambridge University Press, Cambridge, 2009.
[31] J.L. Hodges, E.L. Lehmann, Ranks methods for combination of independent experiments in analysis of variance, Ann. Math. Stat. 33 (1962) 482–497.
[32] S. Holm, A simple sequentially rejective multiple test procedure, Scand. J. Stat. 6 (1979) 65–70.
[33] J. Hühn, E. Hüllermeier, FURIA: an algorithm for unordered fuzzy rule induction, Data Min. Knowl. Discov. 19 (3) (2009) 293–319.
[34] H. Ishibuchi, T. Nakashima, Effect of rule weights in fuzzy rule-based classification systems, Fuzzy Syst. IEEE Trans. 9 (4) (2001) 506–515.
[35] 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.
[36] E.P. Klement, R. Mesiar, E. Pap, Triangular Norms, Kluwer Academic Publisher, Dordrecht, 2000.
[37] E. Lehrer, A new integral for capacities, Econ. Theory 39 (1) (2009) 157–176.
[38] E. Lehrer, R. Teper, The concave integral over large spaces, Fuzzy Sets Syst. 159 (16) (2008) 2130–2144.
[39] G. Lucca, G.P. Dimuro, V. Mattos, B. Bedregal, H. Bustince, J.A. Sanz, A family of Choquet-based non-associative aggregation functions for application in fuzzy rule-based classification systems, in: 2015 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), IEEE, Los Alamitos, 2015.
[40] G. Lucca, J. Sanz, G.P. Dimuro, B. Bedregal, R. Mesiar, A. Kolesárová, H.B. Sola, Pre-aggregation functions: construction and an application, IEEE Trans. Fuzzy Syst. 24 (2) (2016) 260–272.
[41] 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, Knowl. Based Syst. 119 (2017) 32–43.

[42] G. Lucca, J.A. Sanz, G.P. Dimuro, B. Bedregal, R. Mesiar, A. Kolesárová, H. Bustince, The notion of pre-aggregation function, in: V. Torra, T. Narukawa (Eds.), Modeling Decisions for Artificial Intelligence: 12th International Conference, MDAI 2015, Skövde, Sweden, September 21–23, 2015, Proceedings, Springer International Publishing, Cham, 2015, pp. 33–41.
[43] G. Lucca, R. Vargas, G.P. Dimuro, J. Sanz, H. Bustince, B. Bedregal, Analysing some t-norm-based generalizations of the Choquet integral for different fuzzy measures with an application to fuzzy rule-based classification systems, in: P. Santos, R. Prudencio (Eds.), ENIAC 2014 - Encontro Nacional de Inteligência Artificial e Computacional, SBC, São Carlos, 2014.
[44] 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.
[45] R. Mesiar, Choquet-like integrals, J. Math. Anal. Appl. 194 (2) (1995) 477–488.
[46] T. Murofushi, M. Sugeno, Fuzzy t-conorm integral with respect to fuzzy measures: generalization of Sugeno integral and Choquet integral, Fuzzy Sets Syst. 42 (1) (1991) 57–71.
[47] T. Murofushi, M. Sugeno, M. Machida, Non-monotonic fuzzy measures and the Choquet integral, Fuzzy Sets Syst. 64 (1) (1994) 73–86.
[48] R.B. Nelsen, An introduction to copulas, vol. 139 of Lecture Notes in Statistics, Springer, New York, 1999.
[49] J. Sanz, D. Bernardo, F. Herrera, H. Bustince, H. Hagras, A compact evolutionary interval-valued fuzzy rule-based classification system for the modeling and prediction of real-world financial applications with imbalanced data, IEEE Trans. Fuzzy Syst. 23 (4) (2015) 973–990.
[50] J. Sanz, H. Bustince, A. Fernández, F. Herrera, IIVFDT: ignorance functions based interval-valued fuzzy decision tree with genetic tuning, Int. J. Uncertain. Fuzziness Knowl. Based Syst. 20 (supp02) (2012) 1–30.
[51] 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 Trans. Fuzzy Syst. 21 (3) (2013) 399–411.
[52] J.A. Sanz, M. Galar, A. Jurio, A. Brugos, M. Pagola, H. Bustince, Medical diagnosis of cardiovascular diseases using an interval-valued fuzzy rule-based classification system, Appl. Soft Comput. 20 (2014) 103–111.
[53] D. Sheskin, Handbook of Parametric and Nonparametric Statistical Procedures, 2nd, Chapman & Hall/CRC, 2006.
[54] V. Torra, Y. Narukawa, The interpretation of fuzzy integrals and their application to fuzzy systems, Int. J. Approx. Reason. 41 (2006) 43–58.
[55] Z. Wang, G.J. Klir, Generalized Measure Theory, Springer, Boston, 2009.
[56] Z. Wang, K.-S. Leung, M.-L. Wong, J. Fang, A new type of nonlinear integrals and the computational algorithm, Fuzzy Sets Syst. 112 (2) (2000) 223–231.
[57] X. Wen, L. Shao, Y. Xue, W. Fang, A rapid learning algorithm for vehicle classification, Inf. Sci. 295 (2015) 395–406.