Application of the Sugeno Integral in Fuzzy Rule-Based Classification

Jonata Wieczynski[1 [0000−0002−8293−0126]], Giancarlo Lucca[2 [0000−0002−3776−0260]], Eduardo Borges[2 [0000−0003−1595−7676]], and Graçaliz Dimuro[1,2 [0000−0001−6986−9888]]

[1] Departamento de Estadística, Informática y Matemáticas, Universidad Pública de Navarra, Pamplona, Spain
jonata.wieczynski@unavarra.es

[2] Centro de Ciêcias Computacionais, Universidade Federal de Rio Grande, Rio Grande, Brazil
{giancarlo.lucca,eduardoborges,gracalizdimuro}@furg.br

Abstract. Fuzzy Rule-Based Classification System (FRBCS) is a well known technique to deal with classification problems. Recent studies have considered the usage of the Choquet integral and its generalizations to enhance the quality of such systems. Precisely, it was applied to the Fuzzy Reasoning Method (FRM) to aggregate the fired fuzzy rules when classify new data. On the other side, the Sugeno integral, another well known aggregation operator, obtained good results when applied to brain-computer interfaces. Those facts led to the present study in which we consider the Sugeno integral in classification problems. That is, the Sugeno integral is applied in the FRM of a widely used FRBCS and its performance is analyzed over 33 different datasets from the literature. In order to show the efficiency of this new approach, the obtained results are also compared to past studies involving the application of different aggregation functions. Finally, we perform a statistical analysis of the application.

Keywords: Classification problem · Fuzzy Rule-Based Classification System · Fuzzy Reasoning Method · Sugeno integral · Choquet integral.

1 Introduction

Fuzzy Rule-Based Classification Systems (FRBCS's) [18] is a technique used to deal with classification problems [13] which has been applied to diverse problems, e.g., big data [32], image segmentation [21], health [22] and others. The Fuzzy Reasoning Method (FRM) [8, 7] used by FRBCSs is a key component which is composed by four steps. One of them is the aggregation, where the information of the system's fired rules is aggregated, per class. For this step, the FRM normally uses an aggregation function [5], and by doing so, the system will have a different performance (notice that performance in this paper is related to the method accuracy, and not the runtime one) whenever one changes the function.

The work proposed by Barrenechea et al. [4] introduced a new FRM that accounts the usage of all given information by the fired fuzzy rules when classifying a new instance. To do so, they have considered the Choquet integral [10]. Moreover, they introduced a fuzzy measure [30] that is adapted for each class of the problem.

Considering the Choquet integral as basis, was introduced by Lucca et al. [24] the concept of pre-aggregation functions. One way to produce such function is by generalizing the base integral by different t-norms [19]. The generalizations where applied in the FRM to cope with classification problems and elevated the system quality. After that, also considering the Choquet integral as basis, different generalizations where provided and applied, namelY: CC-integral [27], C_F-integral [28], C_{F1F2}-integrals [23] and gC_{F1F2}-integrals [12]. Additionally, this generalizations were also applied in multi-criteria decision making problems [35, 36] and image processing [29][3].

On the otter hand, the Sugeno integral [33] is another fuzzy integral which has been applied to diverse problems in the literature. More recently it was applied to a Motor-Imagery Brain-Computer Interface [20], where it obtained good results when compared to the standard Choquet integral (for more information see [20])
.

Having in consideration that the Choquet integral was used as base to different generalizations and the good results that the Sugeno integral achieved in recent applications, this paper intends to analyze if the usage of the Sugeno integral as aggregation function in the FRM is able to produce a system with competitive results. To do so, we apply and analyze this new base function in the FRM of a state-of-art classifier and provided an analysis over 33 distinct datasets from the literature.

This work is organized as follows. Section 2 presents the background theory in respect to the following sections. In Section 3, the new framework of FRBCS using the Sugeno integral is presented. Then, Section 4 presents and discuss the results. Lastly, Section 5 is the conclusion thoughts of the work.

2 Preliminary concepts and the Sugeno-like generalization

In this section the theoretical background necessary to better understanding of the paper is provided. In what follows consider the following notation: $N = \{1, \ldots, n\}$, that is, the subset of the natural numbers up to n.

An aggregation function (**AF**) [14] is a function $f : [0,1]^n \to [0,1]$ such that the boundary conditions, $f(\mathbf{0}) = 0$ and $f(\mathbf{1}) = 1$, where $\mathbf{0} = (0, \ldots, 0)$ and $\mathbf{1} = (1, \ldots, 1)$, and the monotonicity properties, $\boldsymbol{x} \leq \boldsymbol{y} \implies f(\boldsymbol{x}) \leq f(\boldsymbol{y})$, $\forall \boldsymbol{x}, \boldsymbol{y} \in [0,1]^n$, hold.

A triangular norm (t-norm) is an aggregation function $T : [0,1]^2 \to [0,1]$ that satisfies, for any $x, y, z \in [0,1]$: the commutative ($T(x,y) = T(y,x)$), the

[3] An overview of the different generalizations of the Choquet integral is available in [11].

associative ($T(T(x,y),z) = T(x,T(y,z))$) properties and the boundary condition.

An example of t-norm is the Hamacher t-norm, defined for $x, y \in [0,1]$ as:

$$T_{HP}(x,y) = \begin{cases} 0 & \text{if } \ x = y = 0 \\ \frac{xy}{x+y-xy} & \text{otherwise} \ . \end{cases}$$

A fuzzy measure [33] is a function $m : 2^N \to [0,1]$ that for all $X, Y \subseteq N$ holds the conditions: (i) $m(\emptyset) = 0$ and $m(N) = 1$; (ii) if $X \subset Y$, then $m(X) \leq m(Y)$.

In this study the Power Measure (PM) is considered as the fuzzy measure. It is defined for $X \subseteq N$ as: $m_P(X) = (|X|/n)^q$, with $q > 0$ being genetically learned.

Let m be a fuzzy measure. The standard Choquet integral [6] $\mathfrak{C}_m : [0,1]^n \to [0,1]$ of $\boldsymbol{x} \in [0,1]^n$ with respect to m is defined as:

$$\mathfrak{C}_m(\boldsymbol{x}) = \sum_{i=1}^{n} \left(x_{(i)} - x_{(i-1)}\right) \cdot m(A_{(i)})$$

where (i) is a permutation on 2^N such that $x_{(i-1)} \leq x_{(i)}$ for all $i = 1, \ldots, n$, with $x_{(0)} = 0$ and $A_{(i)} = \{(1), \ldots, (i)\}$.

Let m be a fuzzy measure and $T : [0,1]^2 \to [0,1]$ a t-norm. Then a C_T-integral is defined as $\mathfrak{C}_m^T : [0,1]^n \to [0,1]$, given, for all $\boldsymbol{x} \in [0,1]^n$, by

$$\mathfrak{C}_m^T(\boldsymbol{x}) = \sum_{i=1}^{n} T\left(x_{(i)} - x_{(i-1)}, \ m(A_{(i)})\right)$$

where $x_{(i)}$, $A_{(i)}$ and i is defined as the standard Choquet integral.

Notice that the Choquet integral is an averaging functions [14], i.e., it always holds that for any $\boldsymbol{x} \in [0,1]$ and any fuzzy measure m, $\min(\boldsymbol{x}) \leq \mathfrak{C}_m^T(\boldsymbol{x}) \leq \max(\boldsymbol{x})$

Let Co be a bivariate copula [31]. The Choquet-like integral based on copula with respect to a fuzzy measure m, named CC-integral, is defined as a function $\mathfrak{C}_m^{Co} : [0,1]^n \to [0,1]$, for all $\boldsymbol{x} \in [0,1]^n$, by

$$\mathfrak{C}_m^{Co}(\boldsymbol{x}) = \sum_{i=1}^{n} Co\left(x_{(i)}, \ m(A_{(i)})\right) - Co\left(x_{(i-1)}, \ m(A_{(i)})\right)$$

where $x_{(i)}$, $A_{(i)}$ and i is defined as the standard Choquet integral.

Lastly, the C_F-integral [28] is a generalization of the standard Choquet integral which uses an generic function F instead of the product operator. The definition is as follows: let $F : [0,1]^2 \to [0,1]$ be a function and $m : 2^N \to [0,1]$ a fuzzy measure. Then the C_F-integral $\mathfrak{C}_m^F : [0,1]^n \to [0,1]$ is defined, for all $x \in [0,1]^n$ by:

$$\mathfrak{C}_m^F(\boldsymbol{x}) = \min\left\{1, \sum_{i=1}^{n} F\left(x_{(i)} - x_{(i-1)}, m\left(A_{(i)}\right)\right)\right\},$$

where $x_{(i)}$, $A_{(i)}$ and i is defined as the standard Choquet integral.

In this study as function F the following is considered, $F_{NA} : [0,1]^2 \to [0,1]$:

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

The Sugeno integral is a well know operator, that have been used in many different applications. It is defined with respect to a fuzzy measure m by:

$$Su_m(\boldsymbol{x}) = \bigvee_{i=1}^{n} \left(x_{(i)} \wedge m(A_{(i)}) \right)$$

where $x_{(i)}$, $A_{(i)}$ and i is defined as the Choquet integral. Moreover, it is observable that this integral share the same averaging characteristic as the Choquet integral [14].

3 Application of the Sugeno integral to classification in FRBCS

In this section, the application of the Sugeno integral in a Fuzzy Rule-Based Classification System is presented. We begin presenting the new Fuzzy Reasoning Method that uses of the Sugeno integral. Thereafter, the experimental framework is described. At the end, the obtained results are shown.

3.1 The new Fuzzy Reasoning Method

In this paper, the application of the Sugeno integral take into account a fuzzy classifier that is widely used. Precisely, it considers the Fuzzy Association Rule-based Classification model for High Dimensional Problems (FARC-HD) [1].

The rules used by FARC-HD follows this structure:

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

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 [17], which in this case is computed as the confidence of the fuzzy rule.

Once the fuzzy rules composing the system have been created, the FRM is responsible for classifying new examples. Specifically, let $x_p = (x_{p1}, \ldots, x_{pn})$ be a new example to be classified, L being the number of rules in the rule base, and M being the number of classes of the problem. The new FRM, where the Sugeno integral is used, consist of 4 different steps:

1. To compute the *matching degree*, that is, the strength of the activation of the if-part of the rules for the example x_p, which is computed using a t-norm $T' : [0,1]^n \to [0,1]$:

$$\mu_{A_j}(x_p) = T'(\mu_{A_{j1}}(x_{p1}), \ldots, \mu_{A_{jn}}(x_{pn})), \quad \text{with } j = 1, \ldots, L.$$

2. *Association degree computation*, that is, for the class of each rule the matching degree is weighted with the corresponding rule weight, given by:

$$b_j^k(x_p) = \mu_{A_j}(x_p) \cdot RW_j^k, \quad \text{with } k = Class(R_j), \; j = 1, \ldots, L.$$

3. The *example classification soundness degree for all classes* in this step that the aggregation functions are applied to combine the association degrees obtained in the previous step. The Sugeno integral Su is used as follows:

$$Y_k(x_p) = Su\left(b_1^k(x_p), \ldots, b_L^k(x_p)\right), \quad \text{with } k = 1, \ldots, M. \tag{1}$$

   Since, whenever $b_j^k(x_p) = 0$, it holds that:

$$Su\Big(b_1^k(x_p), \ldots, b_L^k(x_p)\Big) = Su\Big(b_1^k(x_p), \ldots, b_{j-1}^k(x_p), b_{j+1}^k(x_p), \ldots, b_L^k(x_p)\Big),$$

   then, for practical reasons, only those $b_j^k > 0$ are considered in Equation (1).
4. A *Classification* decision function $C : [0,1]^M \to \{1, \ldots, M\}$ is applied over the example classification soundness degrees of all classes and thus, the class corresponding to the maximum soundness degree is determined.

$$C(Y_1, \ldots, Y_M) = \min_{k=1,\ldots,M} k \;\text{ s.t. }\; Y_k = \max_{w=1,\ldots,M}(Y_w).$$

   In practical applications, it is sufficient to consider

$$C(Y_1, \ldots, Y_M) = \arg\max_{k=1,\ldots,M}(Y_k).$$

Finally, its necessary highlight that the fuzzy measure used by the Sugeno and the generalizations of the Choquet integral is the Power Measure, with the exponent q genetically learned as proposed by Barrenechea et al. [4]. This is due to the fact that this fuzzy measure achieved the superior performance in all generalizations. A comparison of the usage of the PM (applied with different generalizations in the FRM) against different fuzzy measure is done in [25].

Table 1: Summary of the datasets used in the study.

Id.	Dataset	#Inst.	#Atts.	#Class	Id.	Dataset	#Inst.	#Atts.	#Class
App	Appendicitis	106	7	2	Pen	Penbased	10,992	16	10
Bal	Balance	625	4	3	Pho	Phoneme	5,404	5	2
Ban	Banana	5,300	2	2	Pim	Pima	768	8	2
Bnd	Bands	365	19	2	Rin	Ring	740	20	2
Bup	Bupa	345	6	2	Sah	Saheart	462	9	2
Cle	Cleveland	297	13	5	Sat	Satimage	6,435	36	7
Con	Contraceptive	1,473	9	3	Seg	Segment	2,310	19	7
Eco	Ecoli	336	7	8	Shu	Shuttle	58,000	9	7
Gla	Glass	214	9	6	Son	Sonar	208	60	2
Hab	Haberman	306	3	2	Spe	Spectfheart	267	44	2
Hay	Hayes-Roth	160	4	3	Tit	Titanic	2,201	3	2
Ion	Ionosphere	351	33	2	Two	Twonorm	740	20	2
Iri	Iris	150	4	3	Veh	Vehicle	846	18	4
Led	led7digit	500	7	10	Win	Wine	178	13	3
Mag	Magic	1,902	10	2	Wis	Wisconsin	683	11	2
New	Newthyroid	215	5	3	Yea	Yeast	1,484	8	10
Pag	Pageblocks	5,472	10	5					

3.2 Experimental Framework

To demonstrate the efficiency and the quality of the proposal, this study uses 33 different datasets. It is necessary to highlight that these datasets are public available in KEEL dataset repository [2]. Also, these datasets are the same used in previous studies (see [28, 23] and [25]).

In Table 1, the characteristics of the datasets are summarized. Then, for each dataset, it is presented the corresponding identification (Id), the number of instances (#Inst), attributes (#Atts), and classes (#Class).

Following the idea of previous generalizations, the results are presented taking into account a 5-fold cross-validation procedure [34]. To analyze the classifier performance the accuracy [34] is used. Consequently, the results presented in this study are related to the average accuracy obtained in the five different folds.

As mentioned before, the fuzzy classifier used in this paper is the FARC-HD, therefore, the configuration of this classifier follows the original author's suggestion. That is, the product t-norm as conjunction operator, the certainty factor is the RW, with 0.05 as minimum support, the threshold for the confidence as 0.8, the depth of the tree is 3, and k_t equals 2.

In relation to the parameters used by the genetic algorithm applied to learn the fuzzy measure, it is considered the same configuration used in different studies ([23],[27] and [28]). To the genetic part of the algorithm it have a population composed by 50 individuals, 30 bits per gene in the gray codification, 20.000 evaluations and the fitness is calculated in therms of the accuracy.

4 Experimental Results

This section describes the obtained results. As discussed in [28] the application of non-averaging functions in the FRM statistically outperformed all the averaging functions. Thus, considering that the proposed FRM, using the Sugeno integral as aggregation, is an averaging approach, in order to provide a fair comparison, in this study we have only performed comparisons against averaging operators.

Again, it is necessary to point out that this study intend to observe that the usage of the Sugeno integral in the FRM can produce a competitive model to deal with classification problems. We are mainly interested in observing if this function is comparable against the standard Choquet integral, since this can allow promising researches on future generalizations of the Sugeno integral, in a similar way that was done with the Choquet integral.

However, aiming at providing a more robust and complete study, comparisons of the new approach against classical FRMs are provided. Precisely, against the Winning Rule (WR) [8], the standard Choquet integral and the best generalizations of the Choquet integral. In this sense we selected the CC-integral (the Choquet integral in its expanded form and generalized by Copulas functions) [27], the best C_T-integral [24] that is based on the Hammacher t-norm and the best averaging C_F-integral that is based in the F_{NA} function.

The obtained results are shown in Table 2. In it, the rows are related to the different datasets (for more details about the dataset see Table 1), per columns different FRMs are compared. The result in each cell is related to the accuracy mean obtained in the cross-validation process. The largest obtained mean in the study, among all approaches, is highlighted in **boldface**.

By taking a general look over the obtained results one can notice that the behavior of FRMs considering the Sugeno integral and the CC-integral are similar. In fact, only in four specific datasets (Ban, Bup, Mag and Two) the achieved result are different.

The biggest obtained accuracy mean is obtained by the C_T- integral, followed closely by the C_F- integral (mean difference of 0.10), CC-integral(mean difference of 0.20) and Sugeno (mean difference of 0.20). Considering the WR and the Choquet integral the obtained mean achieved a low performance.

In a closer look, considering the specific cases where the FRM's provided the largest results (the ones highlighted in **boldface**), the C_T-integral obtained the largest accuracy in 10 of the 33 datasets. However, another interesting result is seen for both, the Sugeno and the CC-integral, where the obtained results are the biggest accuracy in 9 of the 33 datasets. For the remaining cases, the C_F-integral, the Choquet integral and the WR present 6, 4 and 4 of the 33 datasets, respectively. Notice that for the Tit dataset, the obtained means are all equal and therefore are not included in the above count.

By considering and comparing only the Sugeno and Choquet integrals we have that for 19 different datasets the obtained means are superior in favor of the Sugeno integral in comparison to the Choquet (13 cases). On the other hand,

Table 2: Accuracy mean obtained in test by the application of different averaging functions in the FRM.

Dataset	WR	Choquet	CC-integral	C_T-integral	C_F-integral	Sugeno integral
App	83.03	80.13	**85.84**	82.99	82.99	**85.84**
Bal	81.92	82.40	81.60	**82.72**	82.56	81.60
Ban	83.94	**86.32**	84.30	85.96	86.09	85.26
Bnd	69.40	68.56	71.06	**72.13**	69.40	71.06
Bup	62.03	66.96	61.45	65.80	**67.83**	60.87
Cle	56.91	55.58	54.88	55.58	**57.92**	54.88
Com	52.07	51.26	52.61	**53.09**	52.27	52.61
Ecp	75.62	76.51	77.09	**80.07**	78.88	77.09
Gla	64.99	64.02	**69.17**	63.10	64.51	**69.17**
Hab	70.89	72.52	**74.17**	72.21	73.51	**74.17**
Hay	78.69	79.49	**81.74**	79.49	78.72	**81.74**
Ion	90.03	90.04	88.89	89.18	**90.60**	88.89
Iri	**94.00**	91.33	92.67	93.33	93.33	92.67
Led	**69.40**	68.20	68.40	68.60	68.60	68.40
Mag	78.60	78.86	79.81	79.76	**80.02**	79.70
New	94.88	94.88	93.95	**95.35**	93.49	93.95
Pag	94.16	94.16	93.97	**94.34**	93.97	93.97
Pen	**91.45**	90.55	91.27	90.82	**91.45**	91.27
Pho	82.29	82.98	82.94	**83.83**	82.86	82.94
Pim	74.60	73.95	74.21	74.87	**75.64**	74.21
Rin	90.00	**90.95**	87.97	88.78	90.27	87.97
Sah	68.61	69.69	**70.78**	70.77	68.61	**70.78**
Sat	79.63	79.47	79.01	**80.40**	78.54	79.01
Seg	93.03	**93.46**	92.25	93.33	92.55	92.25
Shu	96.00	97.61	**98.16**	97.20	96.78	**98.16**
Son	77.42	77.43	76.95	**79.34**	78.85	76.95
Spe	77.90	77.88	**78.99**	76.02	78.26	**78.99**
Tit	78.87	78.87	78.87	78.87	78.87	78.87
Two	**86.49**	84.46	85.14	85.27	83.92	84.86
Veh	66.67	68.44	**69.86**	68.20	67.97	**69.86**
Win	96.60	93.79	93.83	**96.63**	96.03	93.83
Wis	96.34	**97.22**	95.90	96.78	96.34	95.90
Yea	55.32	55.73	**57.01**	56.53	56.40	**57.01**
Mean	79.15	79.20	79.54	**79.74**	79.64	79.54

when comparing the Sugeno integral against the C_T-integral, the latter achieves superior mean in more than double the number of datasets than the former. However, it is necessary to point that the C_T-integral is a generalization of the Choquet integral, and that the t-norm T was chosen because of its superior results in the FRM, and in this study only the standard Sugeno is considered. Lastly, the results of both F_{NA} and WR are quite similar.

Table 3: Average Rankings of the algorithms by using the Aligned Friedman and the obtained APV

(Pre-)Aggregation function	Ranking	APV
C_T-integral	80.19	
C_F-integral	91.33	0.56
Sugeno integral	97.98	0.56
CC-integral	98.74	0.56
Choquet integral	114.18	<u>0.07</u>
WR	114.56	<u>0.07</u>

Table 4: Results obtained by the Wilcoxon test to pair-wise comparison among the different approaches.

		WR	Choquet	CC-integral	C_T-integral	C_F-integral
	P-value	0.26	0.21	0.87	0.19	0.88
Sugeno integral	R^+	338.5	350.5	250.5	204.5	272
	R^-	222.5	210.5	310.5	356.5	289

4.1 Statistical Analysis

Making comparisons considering the obtained means is a good approach. However, in order to provide a more robust study, in this subsection, we provide a statistical analysis from the different approaches, since it is an interesting question that can enligh the efficiency of the usage of the Sugeno integral.

The statistical analysis consider a non-parametric tests [9], the first analysis is a group comparison using the Aligned Friedman rank test [15]. This test consider a reverse ranking, where the lowest one is considered as control variable and is compared against the others. The results of this test is available in Table 3, which is sorted from the lowest to the largest rank. Also, the Adjusted P-Value (APV) is provided. To calculate the APV the the post-hoc Holm's test [16] is used. In this Table the cases where the null hypothesis is rejected are <u>underlined</u>, having a significance level of 90% ($\alpha = 10\%$).

It can be observed from the group test that the C_T-integral is considered as control method and present statistical differences against the standard Choquet integral and WR. However, when compared to the remaining methods no significant difference were found.

Up to this point, to clarify even more the efficiency of the usage of the Sugeno integral, we have performed a set of pairwise comparisons, with the Wilcoxon signed-rank test [37]. This allows to direct compare the Sugeno integral with the different considered approaches.

The results of the Wilcoxon's test is provided in Table 4. In this table, is shown the obtained p-value, the rank obtained by the Sugeno integral (R^+) and the ranking obtained by the compared method (R^-).

The obtained results reinforce that the Sugeno integral is equivalent to any averaging operator used in different FRMs in the literature, since no statistical difference were found. Moreover, it is observable that comparing our approach against the standard Choquet integral, the obtained ranking is superior.

5 Conclusion

The usage of Fuzzy Rule-Based Classification Systems are an interesting technique to deal with classification problems. The Fuzzy Reasoning Method is the mechanism to perform the classification of different examples. The aggregation used in the FRM is a key point to define the performance of the system.

The usage of the standard Choquet integral in the FRM have been proposed in the literature and provided satisfactory results. After that, many generalizations of this integral where provided, such as: C_T-integral, CC-integral,C_F-integral and otters.

In this paper we provided an application of the Sugeno integral in the FRM. Precisely, the Sugeno integral. This function have been applied, among otters, in the Brain-Computer Interface (BCI) and demonstrated promising results.

In the experimental results we have compared our approach against classical FRMs using the maximum and the Choquet integral and the ones composed by the generalizations of the Choquet integral. The results demonstrated that the Sugeno integral is able to provide superior results in many different datasets and also that this method is statistically equivalent to the compared ones.

Considering the satisfactory obtained results, some future works can be followed. For instance, to create generalizations of the Sugeno integral, e.g. the FG-functional [3], in the FRM and compare the results to past results from generalizations of the Choquet integral. A deep analysis on the characteristics of the datasets (by using data complexity measures for example [26]) that could affect the performance of the classifier by using the Sugeno integral, is another interesting path.

Acknowledgments

The authors would like to thank CNPq (proc. 305805/2021-5, 301618/2019-4), FAPERGS (proc. 19/2551-0001660-3) and Navarra de Servicios y Tecnologías, S.A. (NASERTIC).

References

1. Alcala-Fdez, J., Alcala, R., Herrera, F.: 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), 857–872 (2011)
2. Alcalá-Fdez, J., Sánchez, L., García, S., Jesus, M., Ventura, S., Garrell, J., Otero, J., Romero, C., Bacardit, J., Rivas, V., Fernández, J., Herrera, F.: Keel: a software tool to assess evolutionary algorithms for data mining problems. Soft Computing **13**(3), 307–318 (2009)

3. Bardozzo, F., De La Osa, B., Horanská, L., Fumanal-Idocin, J., delli Priscoli, M., Troiano, L., Tagliaferri, R., Fernandez, J., Bustince, H.: Sugeno integral generalization applied to improve adaptive image binarization. Information Fusion **68**, 37–45 (2021)
4. Barrenechea, E., Bustince, H., Fernandez, J., Paternain, D., Sanz, J.A.: Using the Choquet integral in the fuzzy reasoning method of fuzzy rule-based classification systems. Axioms **2**(2), 208–223 (2013)
5. Beliakov, G., Pradera, A., Calvo, T.: Aggregation Functions: A Guide for Practitioners. Springer, Berlin (2007)
6. Choquet, G.: Theory of capacities. Annales de l'Institut Fourier **5**, 131–295 (1953–1954)
7. Cordón, O., del Jesus, M.J., Herrera, F.: A proposal on reasoning methods in fuzzy rule-based classification systems. International Journal of Approximate Reasoning **20**(1), 21 – 45 (1999)
8. Cordon, O., del Jesus, M.J., Herrera, F.: Analyzing the reasoning mechanisms in fuzzy rule based classification systems. Mathware and Soft Computing **5**(2-3), 321 – 332 (1998)
9. Demšar, J.: Statistical comparisons of classifiers over multiple data sets. Journal of Machine Learning Research **7**, 1–30 (2006)
10. Dias, C.A., Bueno, J.C.S., Borges, E.N., Botelho, S.S.C., Dimuro, G.P., Lucca, G., Fernandéz, J., Bustince, H., Drews Junior, P.L.J.: Using the Choquet integral in the pooling layer in deep learning networks. In: Fuzzy Information Processing. pp. 144–154. Springer International Publishing, Cham (2018)
11. Dimuro, G.P., Fernández, J., Bedregal, B., Mesiar, R., Sanz, J.A., Lucca, G., Bustince, H.: The state-of-art of the generalizations of the Choquet integral: From aggregation and pre-aggregation to ordered directionally monotone functions. Information Fusion **57**, 27 – 43 (2020)
12. Dimuro, G.P., Lucca, G., Bedregal, B., Mesiar, R., Sanz, J.A., Lin, C.T., Bustince, H.: Generalized $C_{F_1F_2}$-integrals: From choquet-like aggregation to ordered directionally monotone functions. Fuzzy Sets and Systems **378**, 44 – 67 (2020)
13. Duda, R.O., Hart, P.E., Stork, D.G.: Pattern Classification (2Nd Edition). Wiley-Interscience (2000)
14. Grabisch, M., Marichal, J.L., Mesiar, R., Pap, E.: Aggregation Functions p. 480 (2009)
15. Hodges, J.L., Lehmann, E.L.: Ranks methods for combination of independent experiments in analysis of variance. Annals of Mathematical Statistics **33**, 482–497 (1962)
16. Holm, S.: A simple sequentially rejective multiple test procedure. Scandinavian Journal of Statistics **6**, 65–70 (1979)
17. Ishibuchi, H., Nakashima, T.: Effect of rule weights in fuzzy rule-based classification systems. Fuzzy Systems, IEEE Transactions on **9**(4), 506–515 (2001)
18. Ishibuchi, H., Nakashima, T., Nii, M.: Classification and Modeling with Linguistic Information Granules, Advanced Approaches to Linguistic Data Mining. Advanced Information Processing, Springer, Berlin (2005)
19. Klement, E.P., Mesiar, R., Pap, E.: Triangular Norms. Kluwer Academic Publisher, Dordrecht (2000)
20. Ko, L., Lu, Y., Bustince, H., Chang, Y., Chang, Y., Fernadez, J., Wang, Y., Sanz, J.A., Pereira Dimuro, G., Lin, C.: Multimodal fuzzy fusion for enhancing the motor-imagery-based brain computer interface. IEEE Computational Intelligence Magazine **14**(1), 96–106 (2019)

21. Leon-Garza, H., Hagras, H., Peña-Rios, A., Conway, A., Owusu, G.: A fuzzy rule-based system using a patch-based approach for semantic segmentation in floor plans. In: 2021 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE). pp. 1–6 (2021)
22. Lixandru-Petre, I.O.: A fuzzy system approach for diabetes classification. In: 2020 International Conference on e-Health and Bioengineering (EHB). pp. 1–4 (2020)
23. Lucca, G., Dimuro, G.P., Fernandez, J., Bustince, H., Bedregal, B., Sanz, J.A.: Improving the performance of fuzzy rule-based classification systems based on a nonaveraging generalization of CC-integrals named $C_{F_1F_2}$-integrals. IEEE Transactions on Fuzzy Systems **27**(1), 124–134 (Jan 2019)
24. Lucca, G., Sanz, J., Pereira Dimuro, G., Bedregal, B., Mesiar, R., Kolesárová, A., Bustince Sola, H.: Pre-aggregation functions: construction and an application. IEEE Transactions on Fuzzy Systems **24**(2), 260–272 (April 2016)
25. Lucca, G., Sanz, J.A., Dimuro, G.P., Borges, E.N., Santos, H., Bustince, H.: Analyzing the performance of different fuzzy measures with generalizations of the choquet integral in classification problems. In: 2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE). pp. 1–6 (June 2019)
26. Lucca, G., Sanz, J., Dimuro, G.P., Bedregal, B., Bustince, H.: Analyzing the Behavior of Aggregation and Pre-aggregation Functions in Fuzzy Rule-Based Classification Systems with Data Complexity Measures, pp. 443–455. Springer International Publishing, Cham (2018)
27. Lucca, G., Sanz, J.A., Dimuro, G.P., Bedregal, B., Asiain, M.J., Elkano, M., Bustince, H.: CC-integrals: Choquet-like copula-based aggregation functions and its application in fuzzy rule-based classification systems. Knowledge-Based Systems **119**, 32 – 43 (2017)
28. Lucca, G., Sanz, J.A., Dimuro, G.P., Bedregal, B., Bustince, H., Mesiar, R.: CF-integrals: A new family of pre-aggregation functions with application to fuzzy rule-based classification systems. Information Sciences **435**, 94 – 110 (2018)
29. Marco-Detchart, C., Lucca, G., Lopez-Molina, C., De Miguel, L., Pereira Dimuro, G., Bustince, H.: Neuro-inspired edge feature fusion using Choquet integrals. Information Sciences **581**, 740–754 (2021)
30. Murofushi, T., Sugeno, M., Machida, M.: Non-monotonic fuzzy measures and the Choquet integral. Fuzzy Sets and Systems **64**(1), 73 – 86 (1994)
31. Nelsen, R.B.: An introduction to copulas. Springer Science & Business Media (2007)
32. da S. E. Tuy, P.G., Nogueira Rios, T.: Summarizer: Fuzzy rule-based classification systems for vertical and horizontal big data. In: 2020 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE). pp. 1–8 (2020)
33. Sugeno, M.: Theory of Fuzzy Integrals and its Applications. Ph.D. thesis, Tokyo Institute of Technology, Tokyo (1974)
34. Tan, P.N., Steinbach, M., Kumar, V.: Introduction to Data Mining, (First Edition). Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA (2005)
35. Wieczynski, J.C., Dimuro, G.P., Borges, E.N., Santos, H.S., Lucca, G., Lourenzutti, R., Bustince, H.: Generalizing the GMC-RTOPSIS method using CT-integral pre-aggregation functions. In: 2020 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE). pp. 1–8. IEEE, Los Alamitos (2020)
36. Wieczynski., J., Lucca., G., Borges., E., Dimuro., G., Lourenzutti., R., Bustince, H.: CC-separation measure applied in business group decision making. In: Proceedings of the 23rd International Conference on Enterprise Information Systems - Volume 1: ICEIS. pp. 452–462. SciTePress (2021)

37. Wilcoxon, F.: Individual comparisons by ranking methods. Biometrics **1**, 80–83 (1945)