

# On the influence of admissible orders in IVOVO\*

Mikel Uriz<sup>1,2</sup>, Daniel Paternain<sup>1,2</sup>[0000-0002-5845-887X], Humberto Bustince<sup>1,2</sup>[0000-0002-1279-6195], and Mikel Galar<sup>1,2</sup>[0000-0003-2865-6549]

<sup>1</sup> Department of Statistics, Computer Science and Mathematics  
Public University of Navarre, Campus Arrosadia s/n, 31006 Pamplona, Spain

<sup>2</sup> Institute of Smart Cities  
Public University of Navarre, Campus Arrosadia s/n, 31006 Pamplona, Spain  
{mikelxabier.uriz,daniel.paternain,bustince,mikel.galar}@unavarra.es

**Abstract.** It is known that when dealing with interval-valued data, there exist problems associated with the non-existence of a total order. In this work we investigate a reformulation of an interval-valued decomposition strategy for multi-class problems called IVOVO, and we analyze the effectiveness of considering different admissible orders in the aggregation phase of IVOVO. We demonstrate that the choice of an appropriate admissible order allows the method to obtain significant differences in terms of accuracy.

**Keywords:** Multi-class classification problems · one-vs-one strategy · interval-valued fuzzy sets · admissible order.

## 1 Introduction

The objective of a classification problem consists in learning a mapping (classifier) from a set of labeled data (examples) into a set of labels (classes), being able of correctly predicting the class of new unseen instances. Depending on the number of classes (labels) the classifier must deal with, the classification problem can either be a two-class (binary) or multi-class problem. From the point of view of complexity, it is much more complex to deal with multi-class problems due to the overlapping between decision boundaries of the classes [16]. Under this context, decomposition strategies, which divide multi-class problems into binary ones, are usually applied to reduce the inherent complexity. The two main decomposition strategies are One-Vs-All (OVA), that generates as many binary problems as number of classes, and One-Vs-One (OVO), that generates as many binary problems as pair of classes, being the latter the most widely used one [10]. Then, each binary problem is solved by an independent base classifier. Finally, each new example is submitted to each classifier and their outputs are fused in order to predict the final class of the example.

---

\* This work has been partially supported by the Spanish Ministry of Science and Technology under the project TIN2016-77356-P and the Public University of Navarre under the project PJUPNA13.

In this work we focus on Fuzzy Rule-Based Classification Systems (FRBCSs) [13], which are known for their interpretability due to the use of linguistic labels in the antecedent of the produced rules. In order to better deal with the uncertainty associated with the choice of fuzzy membership functions in the reasoning method of FRBCSs, in [19] IVTURS, a new method based on the use of interval-valued fuzzy sets, was presented. This model must deal with interval data and therefore, many difficulties arise. One of the most important issue comes from the fact that the usual order between intervals is a partial order.

FRBCSs have been proven to successfully deal with multi-class problems applying an OVO decomposition strategy (see, for example [9, 7, 8]). Moreover, the first solution to handle multi-class problems with IVTURS was given in [8] with the name of IVOVO. It is worth noting that the aggregation phase of IVOVO manages interval-valued confidences and accordingly, the problems related to the order needs to be addressed. The arrangement of interval-valued confidences in IVOVO was solved by the usage of a total order given by Xu and Yager [20], but no further analysis about its suitability has been carried out yet. Moreover, the order given by Xu and Yager has been proven to be a particular case of a larger set of total orders between intervals called *admissible orders* [4]. Considering all these facts, the purpose of this work is to study the influence of different admissible orders (apart from the one given by Xu and Yager) in the performance of IVOVO. To do this, we will consider several admissible orders under two aggregation strategies for OVO, Voting and Win Weighted Voting (WinWV). We will not only evaluate the final performance of each admissible order, but also how influential is on each specific dataset. The experimental study will consider 22 datasets from the KEEL dataset repository [1] and the analysis will be supported by non-parametric statistical tests [11].

The structure of the paper is as follows. In Section 2 we recall the main mathematical concepts related with intervals. In Section 3 we describe FRBCSs, OVO and IVOVO. In Section 4 we explain the concept of admissible order and we provide several construction methods. We finish this paper with an experimental study on the influence of admissible orders in IVOVO in Section 5 and the conclusions and future research lines in Section 6.

## 2 Preliminaries

In this section we recall the theoretical basis of this work. Since we will deal with interval data, we will denote by  $L([0, 1])$  the set of all closed subintervals of the unit interval  $[0, 1]$ , that is,

$$L([0, 1]) = \{\mathbf{x} = [\underline{x}, \bar{x}] \mid (x, \bar{x}) \in [0, 1]^2 \text{ and } \underline{x} \leq \bar{x}\}.$$

*Remark 1.* Although in this work we focus on  $L([0, 1])$ , we can consider also  $L$  as the set of all positive closed intervals, i.e.,  $L = \{\mathbf{x} = [\underline{x}, \bar{x}] \mid 0 \leq \underline{x} \leq \bar{x}\}$ .

Notice that  $L([0, 1])$  is a partially ordered set with respect to the order relation  $\leq_L$  defined in the following way: given  $\mathbf{x}, \mathbf{y} \in L([0, 1])^2$

$$\mathbf{x} \leq_L \mathbf{y} \text{ if and only if } \underline{x} \leq \underline{y} \text{ and } \bar{x} \leq \bar{y}.$$

Considering this order relation, we have that  $(L([0, 1]), \leq_L)$  is a complete lattice whose smallest element is  $\mathbf{0} = [0, 0]$  and the greatest is  $\mathbf{1} = [1, 1]$  [12, 6]

The fact that  $L([0, 1])$  is a partially ordered set means that it is not always possible to establish an order relation between two arbitrary intervals  $\mathbf{x}, \mathbf{y} \in L([0, 1])$ . We will say that  $\mathbf{x}$  and  $\mathbf{y}$  are incomparable, denoted by  $\mathbf{x} \parallel \mathbf{y}$ , whenever  $\mathbf{x} \not\leq_L \mathbf{y}$  and  $\mathbf{y} \not\leq_L \mathbf{x}$  hold simultaneously.

*Example 1.* The intervals  $\mathbf{x} = [0.1, 0.8]$  and  $\mathbf{y} = [0.2, 0.3]$  are incomparable. In fact,  $\mathbf{x} \parallel \mathbf{y}$  whenever  $(\underline{x} < \underline{y} \text{ and } \bar{y} < \bar{x})$  or  $(\underline{y} < \underline{x} \text{ and } \bar{x} < \bar{y})$ .

### 3 IVOVO: Interval-Valued One-Vs-One

IVOVO stands for Interval-Valued One-Vs-One, and is based on the application of the OVO strategy to IVTURS fuzzy classifier, which outputs interval-valued confidence degrees instead of real-valued ones. For this reason, in this section we recall IVOVO and its main components: IVTURS and OVO.

#### 3.1 Fuzzy Rule-Based Classification Systems: IVTURS

Among classification algorithms, Fuzzy Rule-Based Classification Systems (FRBCSs) try to create models whose rules are interpretable by humans due to the use of linguistic labels [13]. These linguistic rules are extracted by a learning algorithm from a training dataset  $\mathcal{D}_T$  having  $P$  labeled examples  $x_p = (x_{p1}, \dots, x_{pn})$ ,  $p = \{1, \dots, P\}$ , where  $x_{pi}$  is the value of the  $i$ -th attribute ( $i = \{1, 2, \dots, n\}$ ) of the  $p$ -th training example. Each example is associated with a class  $y_p \in \mathbb{C} = \{C_1, C_2, \dots, C_m\}$ , being  $m$  is the number of classes of the problem.

IVTURS algorithm [19] is based on FARC-HD (Fuzzy Association Rule-based Classification model for High-Dimensional problems) [1]. Both use rules with the following structure:

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

where  $R_j$  is the label of the  $j$ -th rule,  $x = (x_1, \dots, x_n)$  is a vector representing the example,  $A_{ji} \in \mathbb{X}_i$  is a linguistic label modeled by a triangular membership function (where  $\mathbb{X}_i = \{X_{i1}, \dots, X_{il}\}$  is the set of linguistic labels for the  $i$ -th antecedent, being  $l$  the number of linguistic labels in this set),  $C_j$  is the class label and  $RW_j$  is the rule weight computed using the certainty factor defined in [14].

The main difference in the rule representation between FARC-HD and IVTURS is that the latter take advantage of Interval-Valued Fuzzy Sets (IVFSs) to model the uncertainty under the definition of the linguistic labels, and hence its membership functions are defined by IVFSs instead of FSs. Accordingly, the whole Fuzzy Reasoning Method (FRM) needs to be adapted to work with interval along all its steps. As a consequence, the confidence (association) degree for each class obtained in the final step is also an interval. Therefore, the final class

is taken as the one with the largest confidence degree (according to an admissible order, see Section 4).

With respect to the rule learning algorithm, FARC-HD was composed of three steps (see [1] for more details): a fuzzy association rule extraction, a candidate rule pre-screening, and a genetic rule selection and lateral tuning. IVTURS makes use of FARC-HD for carrying out the rule extraction, but without performing the last step. Then, it introduces IVFSs and finally uses a genetic algorithm to tune the interval FRM and carry out a rule selection.

### 3.2 One-Versus-One (OVO)

In OVO the original  $m$  class problem is transformed into a  $m(m-1)/2$  sub-problems (all possible pair of classes). Therefore, each base classifier will learn to distinguish a pair of classes  $\{C_i, C_j\}$ . To predict the class of a new examples, each classifier is expected to provide a pair confidence degrees  $r_{ij}, r_{ji} \in [0, 1]$  in favor of classes  $C_i$  and  $C_j$ , respectively. For simplicity, these outputs are stored in a *score-matrix*  $R$ . In the case of fuzzy classifiers, these pairs are rarely normalized [9, 7]. This fact requires a normalization step so that the outputs of all the base classifiers are in the same scale. Normalization with real-valued confidence degrees is direct, but it is not so straightforward with intervals.

### 3.3 IVOVO: Interval-Valued One-Vs-One

IVOVO [8] refers to the combination of IVTURS and OVO to enhance the performance of the former in multi-class problems. Nevertheless, there are three main issues when using OVO with IVTURS because the score-matrix is filled by interval confidence scores: 1) there is no consensus on which normalization strategy should be applied; 2) the aggregations needs to be adapted to work with intervals; 3) there is no a total order defined to compare intervals.

Hereafter we recall how these issues were addressed in [8]. Recall that the score-matrix is formed of intervals ( $\mathbf{R}$ ):

$$\mathbf{R} = \begin{pmatrix} - & \mathbf{r}_{12} & \cdots & \mathbf{r}_{1m} \\ \mathbf{r}_{21} & - & \cdots & \mathbf{r}_{2m} \\ \vdots & & & \vdots \\ \mathbf{r}_{m1} & \mathbf{r}_{m2} & \cdots & - \end{pmatrix} \quad (2)$$

$\mathbf{r}_{ij}, \mathbf{r}_{ji} \in L$  corresponding to the confidence degrees for classes  $C_i, C_j$ , respectively.

In IVOVO, the score-matrix  $\mathbf{R}$  was normalized to a new score-matrix  $\mathbf{R}^u$  in such a way that all the elements are closed sub-intervals in  $[0, 1]$ , that is,  $\mathbf{r}_{ij}^u \in L([0, 1])$  for every  $i, j, i \neq j$  (according to the theory described in [19]).

This was done by normalizing them according to the upper bounds:

$$\mathbf{r}_{ij}^u = \begin{cases} \left[ \frac{\underline{r}_{ij}}{\bar{r}_{ij} + \bar{r}_{ji}}, \frac{\bar{r}_{ij}}{\bar{r}_{ij} + \bar{r}_{ji}} \right] & \text{if } \bar{r}_{ij} \neq 0 \text{ or } \bar{r}_{ji} \neq 0 \\ [0.5, 0.5] & \text{otherwise} \end{cases} \quad (3)$$

This normalization allows one to maintain the proportion of ignorance and satisfies the property  $\bar{r}_{ij}^u + \bar{r}_{ji}^u = 1$ .

Regarding the adaptation of the aggregations methods for OVO, they mainly consisted in using the interval arithmetic. We recall the voting strategy and the WinWV strategy as they will be the ones considered in the experimental study (notice that WV was shown to perform worse than WinWV when considering fuzzy classifiers and IVOVO).

- *Voting strategy (Vote):*  $Class = \arg \max_{i=1, \dots, m} \sum_{1 \leq j \neq i \leq m} s_{ij}$ , where  $s_{ij}$  is 1 if  $\mathbf{r}_{ij}^u > \mathbf{r}_{ji}^u$  and 0 otherwise.
- *WinWV:*  $Class = \arg \max_{i=1, \dots, m} \sum_{1 \leq j \neq i \leq m} \mathbf{s}_{ij}$ , where  $\mathbf{s}_{ij}$  is  $\mathbf{r}_{ij}^u$  if  $\mathbf{r}_{ij}^u > \mathbf{r}_{ji}^u$  and 0 otherwise.

*Remark 2.* Observe that in both voting strategies the need to compare intervals appears, thus the necessity of using total orders between intervals.

## 4 Admissible orders for comparing interval data

In the preliminaries of this paper we have seen that the set  $L([0, 1])$  is usually equipped with a partial order  $\leq_L$ . This order, which is not a total order, does not allow neither to compare any arbitrary pair of intervals, nor to calculate the maximum from an arbitrary set of intervals (see, for example [15, 4, 5, 18]). These two aspects are a crucial step of the IVOVO algorithm, specially in the aggregation phase, where intervals have to be compared in both the Voting and the WinWV strategies.

If we focus on the Voting strategy, we observe that each vote is given to the class whose interval-valued confidence is greater. In the original IVOVO, this was done using Xu and Yager's total order. With respect to the WinWV strategy, apart from comparing which interval-valued confidence is greater (in the same way as in the Voting strategy), we must obtain the most voted class, where the vote for each class is again given by an interval. Again, this last calculation was originally performed with Xu and Yager's total order.

In fact, the total order given by Xu and Yager has been widely used in many decision-making procedures both based on interval-valued and intuitionistic fuzzy sets. However, in [4], a wider concept that encompasses Xu and Yager's order, among others, was given establishing therefore the theoretical framework of *admissible orders*. An admissible order is as a linear (total) order that refines the partial order  $\leq_L$ .

**Definition 1.** Consider the partially ordered set  $(L([0, 1]), \leq_L)$ . The order  $\preceq$  on  $L([0, 1])$  is called an admissible order if

- (i)  $\preceq$  is a linear order on  $L([0, 1])$ ;
- (ii) for all  $\mathbf{x}, \mathbf{y} \in L([0, 1])$ ,  $\mathbf{x} \preceq \mathbf{y}$  whenever  $\mathbf{x} \leq_L \mathbf{y}$ .

Consider the following order: we say that  $\mathbf{x} \preceq_{XY} \mathbf{y}$  if and only if  $\underline{x} + \bar{x} \leq \underline{y} + \bar{y}$  or  $(\underline{x} + \bar{x} = \underline{y} + \bar{y} \text{ and } \bar{y} - \underline{y} < \bar{x} - \underline{x})$ . We have that  $\preceq_{XY}$  is an admissible order and it is actually the order given by Xu and Yager.

*Example 2.* Let  $\mathbf{x} = [0, 0.6], \mathbf{y} = [0.2, 0.4], \mathbf{z} = [0.3, 0.3]$ . Observe that any two intervals can be compared by means of  $\leq_L$ . However, according to  $\preceq_{XY}$ , we have that  $\mathbf{x} \preceq_{XY} \mathbf{y} \preceq \mathbf{z}$ , and so  $\min_{\preceq_{XY}} \{\mathbf{x}, \mathbf{y}, \mathbf{z}\} = \mathbf{x}$  and  $\max_{\preceq_{XY}} \{\mathbf{x}, \mathbf{y}, \mathbf{z}\} = \mathbf{z}$ .

Apart from the well-known admissible order given by Xu and Yager, there exist two examples of admissible orders which are derived from the usual lexicographic rules used in  $\mathbb{R}^2$ .

*Example 3.* The following are examples of admissible orders:

- (i)  $\mathbf{x} \preceq_{Lex1} \mathbf{y}$  if and only if  $\underline{x} < \underline{y}$  or  $(\underline{x} = \underline{y} \text{ and } \bar{x} \leq \bar{y})$ .
- (ii)  $\mathbf{x} \preceq_{Lex2} \mathbf{y}$  if and only if  $\bar{x} < \bar{y}$  or  $(\bar{x} = \bar{y} \text{ and } \underline{x} \leq \underline{y})$ .

*Example 4.* Having again  $\mathbf{x} = [0, 0.6], \mathbf{y} = [0.2, 0.4], \mathbf{z} = [0.3, 0.3]$ , we have that  $\mathbf{x} \preceq_{Lex1} \mathbf{y} \preceq_{Lex1} \mathbf{z}$ , while  $\mathbf{z} \preceq_{Lex2} \mathbf{y} \preceq_{Lex2} \mathbf{x}$ .

Notice that the admissible orders given by Lex1 and Lex2 are, in some sense, extreme cases. This means that, for any  $\mathbf{x}, \mathbf{y} \in L([0, 1])$  with  $\mathbf{x} \parallel \mathbf{y}$ , we always have that if  $\mathbf{x} \preceq_{Lex1} \mathbf{y}$ , then  $\mathbf{y} \preceq_{Lex2} \mathbf{x}$ , and viceversa.

Besides these well-known examples of admissible orders, in [4], a construction method of admissible orders considering two continuous functions defined on  $K([0, 1]) = \{(x, y) \in [0, 1]^2 \mid x \leq y\}$  was given.

**Definition 2.** Let  $\preceq$  be an admissible order on  $L([0, 1])$ . The order  $\preceq$  is called a generated admissible order if there exist two continuous functions  $f, g : K([0, 1]) \rightarrow \mathbb{R}$  such that for all  $[a, b], [c, d] \in L([0, 1])$ ,

$$[a, b] \preceq [c, d] \text{ if and only if } [f(a, b), g(a, b)] \preceq_{Lex1} [f(c, d), g(c, d)].$$

We will denote the admissible order generated by the pair  $(f, g)$  as  $\preceq_{f,g}$ .

*Example 5.* Let  $f(x, y) = \frac{x^2 + y^2}{2}$  and  $g(x, y) = y$ . If we consider  $\mathbf{x} = [0, 0.6], \mathbf{y} = [0.2, 0.4], \mathbf{z} = [0.3, 0.3]$ , we have that  $f(\underline{x}, \bar{x}) = 0.18$ ,  $f(\underline{y}, \bar{y}) = 0.1$  and  $f(\underline{z}, \bar{z}) = 0.09$  and, therefore,  $\mathbf{z} \preceq_{f,g} \mathbf{y} \preceq_{f,g} \mathbf{x}$ .

Finally, a much simpler and parametrizable family of admissible orders by considering only two real numbers in the unit interval was also presented in [4]. The admissible order applies the so called  $K_\alpha$  operator, which is a mapping  $K_\alpha : [0, 1]^2 \rightarrow [0, 1]$  given by  $K_\alpha(a, b) = a + \alpha(b - a)$  [3].

**Definition 3.** Let  $\alpha, \beta \in [0, 1]$  with  $\alpha \neq \beta$ . For any  $\mathbf{x}, \mathbf{y} \in L([0, 1])$ , we say that  $\mathbf{x} \preceq_{\alpha, \beta} \mathbf{y}$  if and only if  $K_{\alpha}(\underline{x}, \bar{x}) < K_{\alpha}(\underline{y}, \bar{y})$  or  $(K_{\alpha}(\underline{x}, \bar{x}) = K_{\alpha}(\underline{y}, \bar{y}))$  and  $K_{\beta}(\underline{x}, \bar{x}) < K_{\beta}(\underline{y}, \bar{y})$ .

*Example 6.* Let  $\alpha = 1/3$  and  $\beta = 2/3$ . Then, if we have  $\mathbf{x} = [0, 0.6], \mathbf{y} = [0.2, 0.4], \mathbf{z} = [0.3, 0.3]$ , then  $\mathbf{x} \preceq_{1/3, 2/3} \mathbf{y} \preceq_{1/3, 2/3} \mathbf{z}$ , while  $\mathbf{z} \not\preceq_{2/3, 1/3} \mathbf{y} \preceq_{2/3, 1/3} \mathbf{x}$ .

## 5 Experimental study

The goal of this experimental study is to analyze the influence of admissible orders in the arrangement of interval-valued confidences in both Vote and WinWV strategies. To do so, we will make use of the same experimental framework as in [8] considering a set of 6 different admissible orders, the six admissible orders presented in the examples of Section 4, namely  $\preceq_{Lex1}, \preceq_{Lex2}, \preceq_{1/3, 2/3}, \preceq_{f, g}, \preceq_{2/3, 1/3}$  and  $\preceq_{XY}$ . We recall that any admissible order is a refinement of the partial order  $\leq_L$ , which means that if any pair of intervals can be arranged by means of  $\leq_L$ , then any admissible order  $\preceq$  will produce exactly the same arrangement, i.e., the admissible orders have no influence in the experiment. On the contrary, as the frequency of appearance of incomparable interval-valued confidences increases, the different admissible orders may yield a larger variety of results, thus justifying this analysis.

### 5.1 Experimental framework

The usage of different admissible orders is evaluated using twenty-two datasets from the KEEL dataset repository [2]. Notice that the same datasets were considered in previous studies [9, 7, 8]. Table 1 presents a summary description of the datasets, including the number of examples (#Ex.), the number of attributes (#Atts.), the number of numerical (#Num.) and nominal (#Nom.) attributes, and the number of classes (#Class.).

Table 1: Summary description of the datasets.

Id.	Dataset	#Ex.	#Atts.	#Num.	#Nom.	#Class.	Id.	Dataset	#Ex.	#Atts.	#Num.	#Nom.	#Class.
aut	autos	159	25	15	10	6	bal	balance	625	4	4	0	3
cle	cleveland	297	13	13	0	5	con	contraceptive	1473	9	6	3	3
der	dermatology	358	34	1	33	6	eco	ecoli	336	7	7	0	8
gla	glass	214	9	9	0	7	hay	hayes-roth	132	4	4	0	3
iri	iris	150	4	4	0	3	lym	lymphography	148	18	3	15	4
new	newthyroid	215	5	5	0	3	pag	pageblocks	548	10	10	0	5
pen	penbased	1100	16	16	0	10	sat	satimage	643	36	36	0	7
seg	segment	2310	19	19	0	7	shu	shuttle	2175	9	9	0	7
tae	tae	151	5	3	2	3	thy	thyroid	720	21	21	0	3
veh	vehicle	846	18	18	0	4	vow	vowel	990	13	13	0	11
win	wine	178	13	13	0	3	yea	yeast	1484	8	8	0	10

The results are evaluated using accuracy performance measure obtained by a *stratified 5-fold cross-validation model* following the *Distribution Optimally Balanced Cross Validation* procedure [17]. As recommended in the literature

[11], our conclusions are supported by the proper statistical analysis using non-parametric statistical tests. Wilcoxon rank test is used for pairwise comparisons, whereas Aligned Friedman test is considered for multiple method comparisons.

With respect to the configuration of IVTURS for generating the base classifiers of IVOVO, the configuration recommended by the authors is used: 5 fuzzy labels for each variable, 3 as maximum depth of the tree, a minimum support of 0.05, a minimum confidence of 0.8, 50 individuals as population size, 30 bits per gene for the Gray codification and a maximum of 20000 evaluations.

## 5.2 Influence of normalization strategies in IVOVO

Tables 2 and 3 present the classification accuracy obtained by each admissible order tested using both Vote and WinWV aggregations methods, respectively. Additionally, in each table we include two columns (loss and ad%). Loss refers to the percentage of accuracy loss between the best admissible order and the worst (computed as  $\frac{best-worst}{best}$ ), whereas ad% refers to the percentage of times that the usage of admissible orders can change the interval arrangement (that is, incomparable intervals are present). This information is useful for understanding the possible effect of the admissible order in the loss. Obviously, if ad% is 0, it means that all admissible orders will return the same results and hence, the loss will also be 0.

Table 2: Results VOTE

Dataset	$\preceq_{Lex1}$	$\preceq_{Lex2}$	$\preceq_{1/3,2/3}$	$\preceq_{f,g}$	$\preceq_{2/3,1/3}$	$\preceq_{XY}$	loss	ad%
aut	75.78	72.17	76.52	75.92	75.23	<b>77.13</b>	6.42	1.70
bal	81.10	84.15	84.31	85.27	<b>85.44</b>	85.12	5.08	6.68
cle	<b>56.93</b>	53.54	55.26	55.25	54.58	54.57	5.97	7.07
con	<b>54.65</b>	52.62	53.84	53.23	53.16	53.64	3.72	11.33
der	<b>95.29</b>	<b>95.29</b>	<b>95.29</b>	<b>95.29</b>	<b>95.29</b>	<b>95.29</b>	0.00	0.01
eco	80.18	<b>83.47</b>	82.57	83.17	82.86	81.67	3.94	3.81
gla	70.50	63.40	<b>71.93</b>	69.11	70.55	70.98	11.87	7.78
hay	<b>74.45</b>	<b>74.45</b>	<b>74.45</b>	<b>74.45</b>	<b>74.45</b>	<b>74.45</b>	0.00	0.00
iri	<b>97.33</b>	94.00	96.00	95.33	95.33	95.33	3.42	1.39
lym	80.47	79.83	<b>80.52</b>	<b>80.52</b>	79.83	<b>80.52</b>	0.86	0.14
new	91.63	91.63	<b>95.35</b>	91.16	93.02	94.88	4.39	7.21
pag	89.56	93.81	94.17	<b>94.35</b>	<b>94.35</b>	<b>94.35</b>	5.08	3.65
pen	93.92	90.93	95.10	93.92	94.19	<b>95.19</b>	4.48	6.23
sat	74.52	69.55	79.02	71.69	71.69	<b>81.98</b>	15.16	12.55
seg	90.69	86.71	<b>92.25</b>	90.95	91.34	92.16	6.01	5.69
shu	89.30	82.46	91.87	86.34	86.80	<b>94.35</b>	12.60	6.11
tae	55.53	56.73	56.13	56.13	<b>56.77</b>	<b>56.77</b>	2.19	12.58
thy	93.48	94.03	94.03	94.03	94.03	<b>94.17</b>	0.74	2.92
veh	70.32	61.80	<b>72.21</b>	65.58	68.07	70.91	14.41	17.04
vow	84.04	84.85	89.09	88.18	89.29	<b>89.90</b>	6.52	7.21
win	94.85	95.54	<b>96.63</b>	96.09	96.09	<b>96.63</b>	1.84	3.51
yea	56.53	55.07	<b>59.57</b>	58.43	58.90	<b>59.57</b>	7.55	10.20
AVG	79.59	78.00	81.19	79.75	80.06	<b>81.34</b>	5.56	6.13

Attending at these tables, several points can be highlighted:



Table 3: Results WINWV

Dataset	$\preceq_{Lex1}$	$\preceq_{Lex2}$	$\preceq_{1/3,2/3}$	$\preceq_{f,g}$	$\preceq_{2/3,1/3}$	$\preceq_{XY}$	loss	ad%
aut	75.03	70.90	75.17	75.83	<b>76.44</b>	75.85	7.25	1.70
bal	71.95	83.34	79.96	83.99	<b>84.31</b>	82.54	14.66	6.68
cle	49.47	53.53	55.24	<b>55.93</b>	55.26	54.58	11.54	7.07
con	53.70	52.42	53.70	53.57	53.37	<b>53.71</b>	2.40	11.33
der	<b>96.96</b>	96.69	96.69	96.69	96.69	96.69	0.28	0.01
eco	76.29	<b>84.36</b>	81.35	82.25	83.13	81.37	9.57	3.81
gla	59.90	58.74	62.21	65.01	<b>67.32</b>	64.53	12.75	7.78
hay	<b>75.22</b>	<b>75.22</b>	<b>75.22</b>	<b>75.22</b>	<b>75.22</b>	<b>75.22</b>	0.00	0.00
iri	<b>96.67</b>	94.00	96.00	95.33	95.33	95.33	2.76	1.39
lym	80.49	80.54	<b>81.23</b>	<b>81.23</b>	80.54	<b>81.23</b>	0.91	0.14
new	87.91	90.70	93.02	91.63	93.02	<b>93.95</b>	6.44	7.21
pag	74.11	<b>94.17</b>	83.37	92.10	92.29	88.01	21.30	3.65
pen	88.92	93.28	92.47	94.01	<b>94.56</b>	93.93	5.96	6.23
sat	71.41	66.28	78.71	72.95	73.73	<b>81.83</b>	19.00	12.55
seg	85.54	84.81	88.66	88.57	88.96	<b>89.31</b>	5.04	5.69
shu	67.94	82.37	81.58	89.35	<b>89.99</b>	89.46	24.51	6.11
tae	52.35	56.77	56.19	<b>56.82</b>	<b>56.82</b>	56.79	7.87	12.58
thy	92.51	93.89	93.47	<b>93.90</b>	<b>93.90</b>	93.48	1.48	2.92
veh	65.96	63.22	<b>71.03</b>	67.23	68.89	<b>71.03</b>	10.99	17.04
vow	75.45	82.83	81.31	84.44	<b>85.56</b>	84.24	11.81	7.21
win	93.17	95.54	96.63	96.09	96.09	<b>97.17</b>	4.12	3.51
yea	52.23	55.89	56.74	57.15	57.69	<b>58.03</b>	9.99	10.20
AVG	74.69	77.70	78.64	79.51	<b>79.96</b>	79.92	8.66	6.13

- In both Vote and WinWV, results are affected by the admissible order selected. Interestingly, WinWV is more affected by this decision, leading to worse results than in Vote (e.g.,  $\preceq_{Lex1}$  drops to 74.69% in average with WinWV vs. 79.59% in Vote).
- $\preceq_{XY}$  obtains the highest accuracy in Vote, whereas  $\preceq_{2/3,1/3}$  leads in WinWV.  $\preceq_{2/3,1/3}$  and  $\preceq_{1/3,2/3}$  behave alternatively between Vote and WinWv. Hence, the order seems to be dependent on the OVO aggregation method.
- Observing the maximum loss values we note that deciding the appropriate admissible order is definitely a key factor. In Vote, a 15% of accuracy loss can be obtained if  $\preceq_{Lex2}$  is used instead of  $\preceq_{XY}$ . Similarly, in WinWV this value increases to 24% in shuttle dataset when considering  $\preceq_{Lex1}$  instead of  $\preceq_{2/3,1/3}$ .
- Overall, for IVOVO using either  $\preceq_{Lex1}$  or  $\preceq_{Lex2}$  is harmful (both for Vote and WinWV). This fact is interesting because it tell us that the interval has more information than the one encoded in its bounds. All the other orders are considering values insider the interval and not letting only one bound decide the final order.

Anyway, the previous statements needs to be supported by the appropriate statistical analysis. Table 4 shows the results of the Aligned Friedman Ranks tests, one for each OVO aggregation to focus on the differences among admissible orders.

According to the tests, results are rather different in each OVO aggregation. In Vote,  $\preceq_{1/3,2/3}$  gets the lowest ranks (better), but it does not show significant differences with  $\preceq_{XY}$ . The performance of the other orders is significantly worse.

Table 4: Aligned Friedman test

Method	Vote Rank (p-value)	WinWV Rank (p-value)
$\preceq_{XY}$	41.61 (0.8422)	43.80 (-)
$\preceq_{1/3,2/3}$	39.32 (-)	68.98 (0.0870)
$\preceq_{2/3,1/3}$	67.57 (0.0286+)	44.11 (0.9780)
$\preceq_{f,g}$	73.59 (0.0089+)	51.95 (0.9585)
$\preceq_{Lex1}$	76.02 (0.0058+)	106.70 (0.0000+)
$\preceq_{Lex2}$	100.89 (0.0000+)	83.45 (0.0023+)

+ near the p-value means that statistical differences are found at 95% confidence.

On the contrary, in WinWV  $\preceq_{XY}$  shows its robustness in this case being at the top accompanied by  $\preceq_{2/3,1/3}$  (no significant differences). The rest of the methods are again significantly worse than  $\preceq_{XY}$ . Notice that in both cases,  $\preceq_{Lex1}$  and  $\preceq_{Lex2}$  does not perform well, whereas  $\preceq_{2/3,1/3}$  and  $\preceq_{1/3,2/3}$  have a different behavior depending on the OVO aggregation considered.

Finally, to better understand the results obtained, we tried to find a correlation between the percentage of times the selection of the admissible order may change the decision (ad%) and the percentage of loss obtained. Figure 1 shows these plots for both Vote and WinWV. As we can observe, there is a tendency for greater losses as the ad% increases, which could be expected but should be checked. Pearson correlation coefficients for Vote and WinWV between ad% and loss were 0.6569 and 0.4011, respectively.

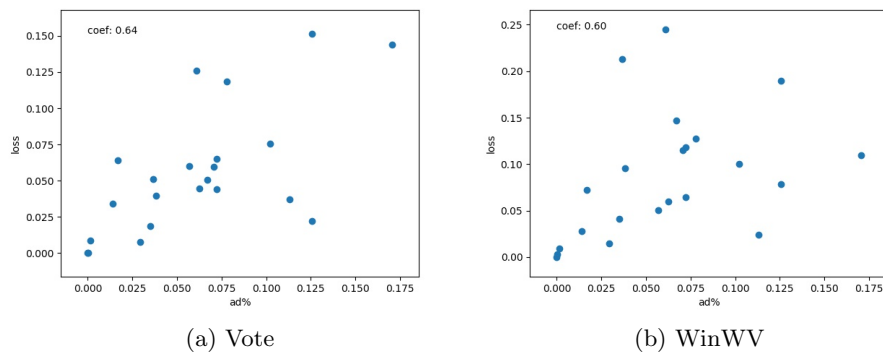


Fig. 1: Scatter plot of ad% vs. loss of all datasets.

## 6 Conclusions

In this work we have analyzed the influence of admissible orders for arranging interval-valued data in IVOVO. To do so, we have considered six different admissible orders, that comes from the most used construction methods, and we have carried out an experimental study that has proven the large differences in terms of accuracy that can be obtained. This result justifies itself the importance of the performed study and points out that deeper studies can be carried out to determine the most appropriate admissible order. From the results obtained, we have seen that the usual order of Xu and Yager can be outperformed by other admissible orders, such as  $\preceq_{2/3,1/3}$  (even  $\preceq_{1/3,2/3}$  obtains very good results). Moreover, we have observed that extreme orders, such as  $\preceq_{Lex1}$  and  $\preceq_{Lex2}$ , are not suitable for IVOVO.

For future work, we aim to carry out a deeper study including more admissible orders and a further analysis of the most suitable order for each aggregation strategy. Moreover, we plan to continue introducing new theoretical developments of interval theory into IVOVO, as interval aggregation functions or interval normalization methods.

## References

1. Alcalá-Fdez, J., Alcalá, R., Herrera, F.: A fuzzy association rule-based classification model for high-dimensional problems with genetic rule selection and lateral tuning. *IEEE Trans. Fuzzy Syst.* **19**(5), 857–872 (2011)
2. Alcalá-Fdez, J., Fernandez, A., Luengo, J., Derrac, J., García, S., Sánchez, L., Herrera, F.: KEEL data-mining software tool: Data set repository, integration of algorithms and experimental analysis framework. *Journal of Multiple-Valued Logic and Soft Computing* **17:2-3**, 255–287 (2011)
3. Bustince, H., Calvo, T., Baets, B.D., Fodor, J., Mesiar, R., Montero, J., Paternain, D., Pradera, A.: A class of aggregation functions encompassing two-dimensional owa operators. *Information Sciences* **180**(10), 1977 – 1989 (2010)
4. Bustince, H., Fernandez, J., Kolesárová, A., Mesiar, R.: Generation of linear orders for intervals by means of aggregation functions. *Fuzzy Sets and Syst.* **220**, 69–77 (2013)
5. Bustince, H., Galar, M., Bedregal, B., Kolesrov, A., Mesiar, R.: A new approach to interval-valued choquet integrals and the problem of ordering in interval-valued fuzzy set applications. *IEEE Transactions on Fuzzy Systems* **21**(6), 1150–1162 (2013)
6. C. Cornelis, G. Deschrijver, E.E.K.: Advances and challenges in interval-valued fuzzy logic. *Fuzzy Sets and Syst.* **157**, 622 – 627 (2016)
7. Elkano, M., Galar, M., Sanz, J., Bustince, H.: 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. *Information Sciences* **332**, 94–114 (2016)
8. Elkano, M., Galar, M., Sanz, J., Lucca, G., Bustince, H.: IVOVO: A new interval-valued one-vs-one approach for multi-class classification problems. In: 17th Int. Fuzzy Sys. Assoc. (IFSA). pp. 1–6 (2017)

9. Elkano, M., Galar, M., Sanz, J.A., Fernández, A., Barrenechea, E., Herrera, F., Bustince, H.: Enhancing multiclass classification in FARC-HD fuzzy classifier: On the synergy between  $n$ -dimensional overlap functions and decomposition strategies. *IEEE Trans. Fuzzy Syst.* **23**(5), 1562–1580 (2015)
10. Galar, M., Fernández, A., Barrenechea, E., Bustince, H., Herrera, F.: An overview of ensemble methods for binary classifiers in multi-class problems: Experimental study on one-vs-one and one-vs-all schemes. *Pattern Recogn.* **44**(8), 1761 – 1776 (2011)
11. García, S., Fernández, A., Luengo, J., Herrera, F.: A study of statistical techniques and performance measures for genetics-based machine learning: Accuracy and interpretability. *Soft Comput.* **13**(10), 959–977 (2009)
12. H. Bustince, J. Montero, M.P.E.B.D.G.: A survey on interval-valued fuzzy sets. *Handbook of Granular Computing*, vol. 1. John Wiley and sons (2008)
13. Ishibuchi, H., Nakashima, T., Nii, M.: Classification and modeling with linguistic information granules: Advanced approaches to linguistic Data Mining. Springer-Verlag (2004)
14. Ishibuchi, H., Yamamoto, T., Nakashima, T.: Hybridization of fuzzy GBML approaches for pattern classification problems. *IEEE Trans. System, Man and Cybernetics B* **35**(2), 359–365 (2005)
15. Lizasoain, I., Moreno, C.: Owa operators defined on complete lattices. *Fuzzy Sets and Systems* **224**, 36 – 52 (2013)
16. Lorena, A., Carvalho, A., Gama, J.: A review on the combination of binary classifiers in multiclass problems. *Artif. Intell. Rev.* **30**(1-4), 19–37 (2008)
17. Moreno-Torres, J., Saez, J., Herrera, F.: Study on the impact of partition-induced dataset shift on k-fold cross-validation. *IEEE Trans. Neural Netw. Learn. Syst.* **23**(8), 1304–1312 (2012)
18. Paternain, D., Miguel, L.D., Ochoa, G., Lizasoain, I., Mesiar, R., Bustince, H.: The interval-valued choquet integral based on admissible permutations. *IEEE Trans. Fuzzy Syst.* (In Press)
19. Sanz, J., Fernández, A., Bustince, H., Herrera, F.: 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), 399–411 (2013)
20. Xu, Z.S., Yager, R.R.: Some geometric aggregation operators based on intuitionistic fuzzy sets. *Int. J. General Syst.* **35**(4), 417–433 (2006)