Application of entropy measures with uncertainty in classification methods with missing data problem

1st Barbara Pękala
University of Rzeszów,
University of Information Technology and Management
Rzeszów, Poland
bpekala@ur.edu.pl

2nd Dawid Kosior
University of Rzeszów
Rzeszów, Poland
dkosior@ur.edu.pl

3rd Krzysztof Dyczkowski
Adam Mickiewicz University
Poznań, Poland
chris@amu.edu.pl

4th Jarosław Szkoła
University of Rzeszów
Rzeszów, Poland
jszkola@ur.edu.pl

***Abstract*—The problem of measuring the degree of entropy based on precedence indicator and similarity measures under conditions of uncertainty or imprecision was studied. So we call back to the notion of precedence and similarity measures of interval-valued fuzzy sets (IVFSs) and we construct an entropy measure with uncertainty by applying for IVFSs of different orders. In addition, we discuss the impact of entropy measures reflecting the uncertainty in the decision-making problem that employed these new measures in the problem of missing values.**

I. Introduction

Representations and investigations in modeling of imprecision and/or uncertainty are still continued since the fuzzy set theory was introduced (Zadeh 1965). One of the many extensions of fuzzy sets (FSs), i.e. interval-valued fuzzy sets (IVFSs) (Zadeh and Sambuc 1975) similar to intuitionistic fuzzy sets (AIFSs) (Atanassov 1986) occurred very useful of their flexibility [1]. Thanks ago this, diverse applications of IVFSs for solving real-life problems involving, for example, pattern recognition, medical diagnosis, or image thresholding were successfully proposed. Among others, many researchers suggested diverse aggregation operators, distances, inclusions, equivalences, similarities, or entropies measures for IVFSs and checked various types of dependencies between them [2] - [11].

The major aspect of our research is to use in the decision-making problem, especially in the problem of missing values, the level of uncertainty represent by entropy measure. To build the entropy measure we use the precedence and similarity between IVFSs reflecting uncertainty. To measure the effectiveness proposed algorithm to find missing values we used a classification problem, the k-nearest neighbors (k-NN) algorithm. In the case of lack of some values in data, a problem of lowering the achievement of classifiers is observed (cf. [12] - [16]). Our aim is to propose a method based on interval measures, i.e. reflects uncertainty holds in many real situations and represents by I-V fuzzy setting. We especially focus on the application of the entropy measure and its construction methods. Entropies are the meaningful aspect of fuzzy sets theory by their high utility [17], [18]. Also, IVFSs [1] in many cases raise the effects of apply fuzzy sets in various applications [19] - [22] for they take into account the uncertainty related to the construction of an exact membership function. Thus, there is widespread interest in studying to extend the concept of entropy to deal with IVFSs [23] - [25]. An essential novelty of the suggested approach is the use of the ordering between intervals and their widths reflecting both uncertainty of data/information and imprecision of the membership functions. Moreover, the adaptation of different types of fusion/aggregation functions was also helpful.

The paper is divided into the following sections. Section 2, basic information of IVFSs given. Inclusion/precedence and similarity degree measures for IVFSs are introduced in Section 3. In Section 4 the construction methods of entropy reflect uncertainty are presented. Finally, an algorithm that uses entropy measures in k-NN classifiers with missing data is examined (Section 5).

II. Interval-valued fuzzy setting

By $L^I = \{[\underline{p}, \overline{p}] : \ \underline{p}, \overline{p} \in [0,1], \ \underline{p} \leqslant \overline{p}\}$ we denoted a family of intervals belonging to the unit interval. If $X \neq \emptyset$, then according to the following papers [26], [17], [27] and [28] we define an **interval-valued fuzzy set** (IVFS) S in X as a mapping $S : X \to L^I$ such that for each $x \in X$

$$S(x) = [\underline{S}(x), \overline{S}(x)]$$

means the degree of membership of an element x into S. The family of all IVFSs in X we denoted by $\mathrm{IVFS}(X)$. We assume, reflect an aspect of applications on a finite set $X = \{x_1, \ldots, x_n\}$. In opposite to fuzzy sets, in the case of the IVFSs, the membership of an element x is not exactly indicated. We only specify an upper and lower bound of the possible membership. This is cause the IVFSs spend so useful for the uncertainty of information. Of course, each fuzzy set S could be treated as the IVFS such that $\underline{S}(x) = \overline{S}(x) \ \forall x \in X$. Thus, $\mathrm{FS}(X) \subset \mathrm{IVFS}(X)$, where $\mathrm{FS}(X)$ is a family of fuzzy sets on X.

The basic operations in IVFSs for $x \in X$ called intersection, union, complement are:

$$S \cap T = \{\langle x, [\min\{\underline{S}(x), \underline{T}(x)\}, \min\{\overline{S}(x), \overline{T}(x)\}]\rangle\}, \quad (1)$$

$$S \cup T = \{\langle x, [\max\{\underline{S}(x), \underline{T}(x)\}, \max\{\overline{S}(x), \overline{T}(x)\}]\rangle\}, \quad (2)$$

$$S_N = \{\langle x, [N(\overline{S}(x)), N(\underline{S}(x))]\rangle\}, \quad (3)$$

978-1-6654-4407-1/21/$31.00 © 2021 European Union

where N is fuzzy negation and particularly, for the standard negation we have
$S^c = \{\langle x, [1-\overline{S}(x), 1-\underline{S}(x)]\rangle\}$
make up a distributive lattice $(\mathrm{IVFS}(X), \cap, \cup)$ and fulfill the Morgan's laws.

Before initiating issues of entropy, inclusion/precedence, or similarity indicators we need to look into ordering relations and aggregations in L^I.

A. Orders in the interval setting

In this section, we concentrate on relations between arbitrary IVFSs $S(x)$ and $T(x)$ for any fixed $x \in X$, so let us assume the following notation
$S(x) = [\underline{S}(x), \overline{S}(x)] = [\underline{s}, \overline{s}]$, $T(x) = [\underline{T}(x), \overline{T}(x)] = [\underline{t}, \overline{t}]$.

The best known and often used partial order in L^I it is

$$[\underline{s}, \overline{s}] \leqslant_2 [\underline{t}, \overline{t}] \Leftrightarrow \underline{s} \leqslant \underline{t} \quad \text{and} \quad \overline{s} \leqslant \overline{t}, \tag{4}$$

where $[\underline{s}, \overline{s}] <_2 [\underline{t}, \overline{t}]$ if and only if $[\underline{s}, \overline{s}] \leqslant_2 [\underline{t}, \overline{t}]$ and ($\underline{s} < \underline{t}$ or $\overline{s} < \overline{t}$).

The fundamental operations, i.e joint and meet are determined in L^I as follows

$$[\underline{s}, \overline{s}] \vee [\underline{t}, \overline{t}] = [\max(\underline{s}, \underline{t}), \max(\overline{s}, \overline{t})],$$
$$[\underline{s}, \overline{s}] \wedge [\underline{t}, \overline{t}] = [\min(\underline{s}, \underline{t}), \min(\overline{s}, \overline{t})].$$

Analogously to $(\mathrm{IVFS}(X), \cap, \cup)$ the structure $(L^I, \vee, \wedge)$ is a lattice with order $\leqslant_2$, where $1_{L^I} = [1, 1]$ and $0_{L^I} = [0, 0]$ are the top and the bottom of $(L^I, \leqslant_2)$, respectively.

In real-life problems we need often to be able to compared data with uncertainty, e.g. intervals, use some a linear order. So we must extended the partial order $\leqslant_2$ to a linear one. Thus, the following relation, so-called, admissible order we recall.

Definition 1 ([29], Def. 3.1). An order $\leq_{Adm}$ in L^I is called **admissible** if $\leq_{Adm}$ is linear in L^I and for all $p, q \in L^I$ $p \leq_{Adm} q$ whenever $p \leqslant_2 q$.

Admissible order has been widely researched, e.g. in [30]. Now we recall the construction of admissible linear order build by aggregation functions presented in [29].

Proposition 1 ([29], Prop. 3.2). *Let $\psi, \upsilon : [0,1]^2 \to [0,1]$ be continuous aggregation functions, which, for all $p = [\underline{p}, \overline{p}], q = [\underline{q}, \overline{q}] \in L^I$, the equalities $\psi(\underline{p}, \overline{p}) = \psi(\underline{q}, \overline{q})$ and $\upsilon(\underline{p}, \overline{p}) = \upsilon(\underline{q}, \overline{q})$ hold iff $p = q$. If the order $\leqslant_{\psi,\upsilon}$ on L^I is defined by*

$$p \leqslant_{\psi,\upsilon} q \Leftrightarrow \psi(\underline{p}, \overline{p}) < \psi(\underline{q}, \overline{q})$$
$$\text{or} \ \ (\psi(\underline{p}, \overline{p}) = \psi(\underline{q}, \overline{q}) \text{ and } \upsilon(\underline{p}, \overline{p}) \leqslant \upsilon(\underline{q}, \overline{q})), \tag{5}$$

then $\leqslant_{\psi,\upsilon}$ is an admissible order in L^I.

In the following, the notation $<_{\psi,\upsilon}$ will indicate that in the strict inequality holds in (5).

Basic examples of admissible orders in L^I ([29]) are:

- the Xu-Yager order [31]

$$[\underline{p}, \overline{p}] \leqslant_{XY} [\underline{q}, \overline{q}] \Leftrightarrow \underline{p} + \overline{p} < \underline{q} + \overline{q}$$
$$\text{or} \ \ (\overline{p} + \underline{p} = \overline{q} + \underline{q} \text{ and } \overline{p} - \underline{p} \leqslant \overline{q} - \underline{q})$$

- lexicographical orders

$$[\underline{p}, \overline{p}] \leqslant_{Lex1} [\underline{q}, \overline{q}] \Leftrightarrow \underline{p} < \underline{q} \ \text{ or } \ (\underline{p} = \underline{q} \text{ and } \overline{p} \leqslant \overline{q}) \tag{6}$$
$$[\underline{p}, \overline{p}] \leqslant_{Lex2} [\underline{q}, \overline{q}] \Leftrightarrow \overline{p} < \overline{q} \ \text{ or } \ (\overline{p} = \overline{q} \text{ and } \underline{p} \leqslant \underline{q}) \tag{7}$$

- the $\alpha\beta$ order

$$p \leqslant_{\alpha\beta} q \Leftrightarrow K_\alpha(\underline{p}, \overline{p}) < K_\alpha(\underline{q}, \overline{q}) \tag{8}$$
$$\text{or} \ \ (K_\alpha(\underline{p}, \overline{p}) = K_\alpha(\underline{q}, \overline{q}) \text{ and } K_\beta(\underline{p}, \overline{p}) \leqslant K_\beta(\underline{q}, \overline{q})),$$

where $K_\alpha : [0,1]^2 \to [0,1]$ is defined as $K_\alpha(a,b) = \alpha a + (1-\alpha)b$ for some $\alpha, \beta \in [0,1]$, $\alpha \neq \beta$ and $a, b \in L^I$.

Remark 1. The notation $<_{Adm}$ will denote, in this paper, that in (5) the strict inequality holds.

B. Interval-valued (I-V) aggregation functions

Now we revise the notion of an aggregation function on L^I being a significant concept in numerous applications. What follows is the description of aggregation functions connected with $\leqslant_2$ and $\leqslant_{Adm}$.

Remark 2. What is emphasized in the present paper is the use of the concept $\leq$ for the either partial or admissible linear order, with 0_{L^I} and 1_{L^I} as marginal and maximum element of L^I, correspondingly. Turning to the results for the partial order, the formerly discussed representation $\leq_2$ is to be applied in a situation in which the results for the admissible linear orders include notation $\leq_{Adm}$ (occasionally including the appropriate instance of such admissible linear order).

Definition 2 ([30], [32], [33]). Let $n \in \mathbf{N}$, $n \geqslant 2$. An operation $\mathcal{A} : (L^I)^n \to L^I$ is called an interval-valued (I-V) aggregation function if it is increasing with regard to the order $\leq$ (partial or linear (see Remark 2)), i.e.

$$\forall x_i, y_i \in L^I \quad x_i \leq y_i \Rightarrow \mathcal{A}(x_1, ..., x_n) \leq \mathcal{A}(y_1, ..., y_n) \tag{9}$$
$$\text{and} \ \ \mathcal{A}(\underbrace{0_{L^I}, ..., 0_{L^I}}_{n\times}) = 0_{L^I}, \ \ \mathcal{A}(\underbrace{1_{L^I}, ..., 1_{L^I}}_{n\times}) = 1_{L^I}.$$

The particular case of operation of I-V aggregations is a representable I-V aggregation function with regard to $\leq_2$.

Definition 3 ([34], [35]). The I-V aggregation function $\mathcal{A} : (L^I)^n \to L^I$ is coined representable in a situation when there exist aggregation functions $A_1, A_2 : [0,1]^n \to [0,1]$ as follows

$$\mathbf{A}(x_1, ..., x_n) = [A_1(\underline{x}_1, ...\underline{x}_n), A_2(\overline{x}_1, ...\overline{x}_n)]$$

for all $x_1, ..., x_n \in L^I$.

The following result offer the characterization of representable I-V aggregation functions on L^I.

Theorem 1 ([36]). *The $\mathcal{A} : (L^I)^n \to L^I$ operation is a representable I-V aggregation function in connection to $\leq_2$ if and only if there are two aggregation functions $A_1, A_2 : [0,1]^n \to [0,1]$ which fulfill for all $x_1, ..., x_n \in L^I$, $A_1 \leqslant A_2$ as well as*

$$\mathcal{A}(x_1, ..., x_n) = [A_1(\underline{x}_1, ...\underline{x}_n), A_2(\overline{x}_1, ...\overline{x}_n)]. \tag{10}$$

Example 1. The fundamental cases of representable I-V aggregation functions on L^I include two operations $\wedge$ as well as $\vee$ on L^I with $A_1 = A_2 = \min$ and $A_1 = A_2 = \max$, respectively (as for the order $\leq_2$, however not $\leq_{Lex1}$, $\leq_{Lex2}$ or $\leq_{XY}$).
The other examples of representable I-V aggregation functions concerning $\leq_2$ are the following:

- the representable arithmetic mean $\mathcal{A}_{mean}([\underline{x},\overline{x}],[\underline{y},\overline{y}]) = [\frac{\underline{x}+\underline{y}}{2}, \frac{\overline{x}+\overline{y}}{2}]$,
- the representable geometric mean $\mathcal{A}_{gmean}([\underline{x},\overline{x}],[\underline{y},\overline{y}]) = [\sqrt{\underline{xy}}, \sqrt{\overline{xy}}]$,
- the representable mean-power mean $\mathcal{A}_{meanpow}([\underline{x},\overline{x}],[\underline{y},\overline{y}]) = [\frac{\underline{x}+\underline{y}}{2}, \sqrt{\frac{\overline{x}^2+\overline{y}^2}{2}}]$,
- the representable product $\mathcal{A}_{prod}([\underline{x},\overline{x}],[\underline{y},\overline{y}]) = [\underline{xy}, \overline{xy}]$,
- the representable prod-mean $\mathcal{A}_{prodmean}([\underline{x},\overline{x}],[\underline{y},\overline{y}]) = [\underline{xy}, \frac{\overline{x}+\overline{y}}{2}]$,
- the representable mean-max $\mathcal{A}_{meanmax}([\underline{x},\overline{x}],[\underline{y},\overline{y}]) = [\frac{\underline{x}+\underline{y}}{2}, \max(\overline{x},\overline{y})]$.

Moreover, an operation

$$\mathcal{A}_\alpha(x,y) = [\alpha\underline{x} + (1-\alpha)\underline{y}, \alpha\overline{x} + (1-\alpha)\overline{y}]$$

is an I-V aggregation function on L^I with regard to $\leq_{Lex1}$, $\leq_{Lex2}$ and $\leq_{XY}$ (see [30]) for $x,y \in L^I$.

III. Inclusion and similarity degree measures for IVFSs

There are two ways in which the inclusion measures, also known as subsethood measures, have been analysed, these are constructive approaches and axiomatic approaches. The inclusion measure was effectively introduced to the theory of extensions of fuzzy sets.

An inclusion measure of fuzzy set A in B for the two fuzzy sets A and B is being characterised as a subset of A in B. Numerous authors tried to provide generalization of the inclusion measure definition, which is *including degree*, to represent and measure the uncertainty data. Some indicators were put forward to specify the degree to which an IVFS is a subset of different IVFS ([30], [21], [37], [38]). We will characterise the inclusion grade indicator of set S in set T for every $S,T \in IVFS(X)$, by the measure of inclusion between their elements, so called intervals. This has led to establishing next considerations on the topic of inclusion measures in an interval settings.

A. Precedence indicator

The another tool useful for manipulations intervals must be introduced before we are able to present a perspective regarding similarity and entropy in the IVFSs environment.

We use the following concept of an inclusion measure employing either linear or partial order as well as uncertainty measure/width of intervals: $w(p) = \overline{p} - \underline{p}$ refers to the width of $p \in L^I$.

We study a precedence indicator concept in which a strong inequality between inputs leads to the same value (the biggest) of the inclusion measure for these inputs. Moreover, in a width of the interval is used in reflexivity property.

Furthermore, the authors recall that $<$ and $\leq$ fulfill Remark 1 and 2.

Definition 4 ([39], cf. [41]). The function $\mathrm{Prec} : (L^I)^2 \to L^I$ is perceived to be a **precedence indicator** if it fulfills the following conditions for any $p,q,u \in L^I$:
P1 if $p = 1_{L^I}$ and $q = 0_{L^I}$ then $\mathrm{Prec}(p,q) = 0_{L^I}$;
P2 if $p < q$, then $\mathrm{Prec}(p,q) = 1_{L^I}$ for any $p,q \in L^I$;
P3 $\mathrm{Prec}(p,p) = [1-w(p), 1]$ for any $p \in L^I$;
P4 if $p \leq q \leq u$, then $\mathrm{Prec}(u,p) \leq \mathrm{Prec}(q,p)$ and $\mathrm{Prec}(u,p) \leq \mathrm{Prec}(u,q)$, for any $p,q,u \in L^I$.

Remark 3. If $p = q$ and $w(p) = 0$, then $\mathrm{Prec}(p,q) = 1_{L^I}$.

Below are presented instances of the constructions of precedence indicator satisfying Definition 4.

Proposition 2 ([41]). *For $p,q \in L^I$ the operation $\mathrm{Prec}_{\mathcal{A}} : (L^I)^2 \to L^I$ is the precedence indicator*

$$\mathrm{Prec}_{\mathcal{A}}(p,q) = \begin{cases} [1-w(p), 1], & p = q, \\ 1_{L^I}, & p < q, \\ \mathcal{A}(N_{IV}(p), q), & \textit{otherwise} \end{cases}$$

for $p,q \in L^I$ and interval-valued (I-V) fuzzy negation N_{IV} (antitonic operation that satisfies $N_{IV}(0_{L^I}) = 1_{L^I}$ and $N_{IV}(1_{L^I}) = 0_{L^I}$, cf. [19], [42]), such that $N_{IV}(p) = [n(\overline{p}), n(\underline{p})] \leq [1-\overline{p}, 1-\underline{p}]$, where n is a fuzzy negation and $\mathcal{A}$ is a representable I-V aggregation such that $\mathcal{A} \leq \vee$.

Example 2. The following function is the precedence indicator with respect to $\leqslant_2$:

$$\mathrm{Prec}_{\mathcal{A}_{meanL^I}}(x,y) = \begin{cases} [1-w(x), 1], & x = y, \\ 1_{L^I}, & x <_2 y, \\ [\frac{1-\overline{x}+\underline{y}}{2}, \frac{1-\underline{x}+\overline{y}}{2}], & \text{otherwise}, \end{cases}$$

$$\mathrm{Prec}_{\mathcal{A}_{meanpow}}(x,y) = \begin{cases} [1-w(x), 1], & x = y, \\ 1_{L^I}, & x <_2 y, \\ [\frac{1-\overline{x}+\underline{y}}{2}, \sqrt{\frac{(1-\underline{x})^2+\overline{y}^2}{2}}], & \text{otherwise}, \end{cases}$$

$$\mathrm{Prec}_{\mathcal{A}_{meanmax}}(x,y) = \begin{cases} [1-w(x), 1], & x = y, \\ 1_{L^I}, & x <_2 y, \\ [\frac{1-\overline{x}+\underline{y}}{2}, \max(1-\underline{x}, \overline{y})], & \text{otherwise}, \end{cases}$$

where $N_{IV}(x) = [1-\overline{x}, 1-\underline{x}]$. Moreover, the function below is the precedence indicator in connection to $\leq_{Lex2}$

$$\mathrm{Prec}_{\mathcal{A}_{meanLex2}}(x,y) = \begin{cases} [1-w(x), 1], & x = y, \\ 1_{L^I}, & x <_{Lex2} y, \\ [\frac{\underline{y}}{2}, \frac{1-\overline{x}+\overline{y}}{2}], & \text{otherwise}, \end{cases}$$

and using the I-V aggregation function $\mathcal{A}_\alpha$ for $\alpha \in [0,1]$, we get the precedence indicator

$$\mathrm{Prec}_{\mathcal{A}_\alpha Lex2}(x,y) = \begin{cases} [1-w(x), 1], & x = y, \\ 1_{L^I}, & x <_{Lex2} y, \\ [(1-\alpha)\underline{y},\ \alpha(1-\overline{x}) + (1-\alpha)\overline{y}], & \text{otherwise}, \end{cases}$$

where

$$N_{IV}(x) = \begin{cases} 1_{L^I}, & x = \mathbf{0}, \\ [0, 1-\overline{x}], & \text{otherwise} \end{cases}$$

is the I-V fuzzy negation in connection to $\leq_{Lex2}$.

B. Similarity measure

In this section, we will recall the class of similarity measures between IVFSs. The impetus for this approach was, firstly, that we were expanding all the concepts of the total order of the intervals, and secondly, that we were considering the width of the intervals in such a way that the uncertainty of the output data is strongly related to the uncertainty of the input data. To construct interval-valued (I-V) similarity, we need I-V aggregation functions and precedence indicators that take into account the width of the intervals.

Let $X \neq \emptyset$ and $card(X) = n$. For $S, T \in IVFS(X)$ and $card(X) = n, n \in N$ we will use the following notion of partial order

$$S \preceq T \Leftrightarrow s_i \leq t_i$$

for $i = 1, ..., n$, where $\leq$ is the same kind of orders (partial or linear, see Remark 1 and 2) for each i and $s_i = S(x_i)$, $t_i = T(x_i)$. Let us note that if for $i = 1, ..., n$ we take the same linear order $s_i \leq t_i$, then the order $S \preceq T$ between IVFSs S, T is partial, but need not be linear.

Definition 5 ([39], cf. [40])**.** Let $A_1 : [0,1]^n \to [0,1]$ be an aggregation function. Then action $S^M : IVFS(X) \times IVFS(X) \to L^I$, which meets the conditions:

(S_{v1}) $S^M(S,T) = S^M(T,S)$ for $S, T \in IVFS(X)$;
(S_{v2}) $S^M(S,S) = [1 - A_1(w_S(x_1), ..., w_S(x_n)), 1]$;
(S_{v3}) $S^M(S,T) = 0_{L^I}$, if $\{S(x_i), T(x_i)\} = \{0_{L^I}, 1_{L^I}\}$;
(S_{v4}) if $S \preceq T \preceq U$, then $S^M(S,U) \leq S^M(S,T)$ and $S^M(S,U) \leq S^M(T,U)$

is called a similarity measure for $i = 1, ..., n$..

Proposition 3 ([39])**.** *Let* Prec *be a precedence indicator. If $\mathcal{A} = [A_1, A_2]$, $\mathcal{B} = [B_1, B_2]$ are representable I-V aggregation functions for which A_1 is as in Def. 5 and self-dual, $\mathcal{B}$ is symmetric with the neutral element 1_{L^I} and B_1 is idempotent aggregation function, then the action $S^M : IVFS(X) \times IVFS(X) \to L^I$:*

$$S^M(S,T) = \mathcal{A}_{i=1}^n(\mathcal{B}(\mathrm{Prec}(S(x_i), T(x_i)), \mathrm{Prec}(T(x_i), S(x_i))))$$

is a similarity measure.

The following example presents direct conclusions from the above theorem.

Example 3. The action $S^M : IVFS(X) \times IVFS(X) \to L^I$:

$$S^M(S,T) = \mathcal{A}_{i=1}^n(\mathrm{Prec}_{\mathcal{A}}(S(x_i), T(x_i) \wedge \mathrm{Prec}_{\mathcal{A}}(T(x_i), S(x_i))))$$

is a similarity measure, where $\mathcal{A} \in \{\mathcal{A}_{mean}, \mathcal{A}_{meanpow}, \mathcal{A}_{meanmax}\}$ with adequate precedence indicators: $\mathrm{Prec}_{\mathcal{A}_{mean}}$, $\mathrm{Prec}_{\mathcal{A}_{meanpow}}$ and $\mathrm{Prec}_{\mathcal{A}_{meanmax}}$.

IV. Entropy measure

We will now present a class of entropy measures for IVFSs. To construct the new interval-valued (I-V) entropy, we need I-V aggregation functions and I-V inclusion measure in connection to the width of the intervals. For $p \in L^I$ we denote by $\mathbf{P} \in IVFS(X)$ such that $\mathbf{P}(x) = p$ for all $x \in X$. Inspiring by [18] by put the following definition of entropy for partial or linear order.

Definition 6. Let N be a strong (involutive) I-V negation with equilibrium point $e \in L^I$. A function $E : IVFS(X) \to L^I$ is an interval-valued (I-V) entropy on IVFS(X) with respect to the negation N if for $S, T \in IVFS(X)$:

(E1) $E(S) = 0_{L^I}$ iff S is crisp;
(E2) $E(\mathbf{E}) = [1 - w(e), 1]$;
(E3) $E(S) \leq E(T)$, if $S(x) \leq T(x) \leq e$ or $S(x) \geq T(x) \geq e$ for all $x \in X$.

The method of the construction of the I-V entropy in terms of a subsethood measure is presented below.

Proposition 4. *Let* Prec *be subsethood measure, N be strong (involutive) I-V negation with equilibrium point e and $\mathcal{A}$ be I-V idempotent aggregation function. Then, the function $E : IVFS(X) \to L^I$:*

$$E(S) = \mathcal{A}_{i=1}^n \mathrm{Prec}_{\mathcal{A}}(S(x_i) \vee S_N(x_i), N(S(x_i) \vee S_N(x_i))) \tag{11}$$

is the I-V entropy with respect to N.

Proof. If S is a crisp, then $(S \vee S_N)(x_i) = 1_{L^I}$ for each i, so $\mathrm{Prec}_{\mathcal{A}}(S(x_i) \vee S_N(x_i), N(S(x_i) \vee S_N(x_i))) = 0_{L^I}$ for each i and as consequence we obtain with property of aggregation function $E(S) = 0_{L^I}$ and (E1) is proved.
(E2) we may obtain directly by P3 and idempotency of aggregation function $\mathcal{A}$.
If $S \preceq T \preceq \mathbf{E}$, then $S \preceq T \preceq \mathbf{E} \preceq T_N \preceq S_N$ and as a consequence of P4 we have

$$E(S) = \mathcal{A}_{i=1}^n \mathrm{Prec}_{\mathcal{A}}(S(x_i) \vee S_N(x_i), N(S(x_i) \vee S_N(x_i))) =$$
$$\mathcal{A}_{i=1}^n \mathrm{Prec}_{\mathcal{A}}(S_N(x_i), S(x_i)) \leq \mathcal{A}_{i=1}^n \mathrm{Prec}_{\mathcal{A}}(T_N(x_i), S(x_i))$$
$$\leq \mathcal{A}_{i=1}^n \mathrm{Prec}_{\mathcal{A}}(T_N(x_i), T(x_i)) = E(T)$$

for $i = 1, ..., n$.
In similar way we may prove case $S \succeq T \succeq \mathbf{E}$. What finished the proof. $\square$

Example 4. By $\mathcal{A}_{mean}$ I-V aggregation function, the precedence indicator

$$\mathrm{Prec}_{\mathcal{A}} \in \{\mathrm{Prec}_{\mathcal{A}_{meanL^I}}, \mathrm{Prec}_{\mathcal{A}_{meanpow}}, \mathrm{Prec}_{\mathcal{A}_{meanmax}}\}$$

and I-V negation $N(x) = [1 - \overline{x}, 1 - \underline{x}]$ we obtain entropy satisfy the equation (11) compared to the standard I-V negation N (see Proposition 2).

The second method construction of the I-V entropy based on the similarity measure.

Proposition 5. *Let S^M be similarity measure satisfy Proposition 3, where A_1 be idempotent aggregation function. Then, the function $E : IVFS(X) \to L^I$:*

$$E(S) = S^M(S, S_N) \tag{12}$$

is the I-V entropy with respect to N (involutive I-V negation with equilibrium point e).

Proof. If S is a crisp, then

$$\{\mathrm{Prec}(S(x_i), T(x_i)), \mathrm{Prec}(T(x_i), S(x_i))\} \in \{0_{L^I}, 1_{L^I}\}$$

and because $\mathcal{B}$ has neutral element 1_{L^I} we obtain $E(S) = 0_{L^I}$.
(E2) we may obtain directly by (Sv2) and idempotency of aggregation function $\mathcal{A}$.
If $S \preceq T \preceq \mathbf{E}$, then $S \preceq T \preceq \mathbf{E} \preceq T_N \preceq S_N$ and as a consequence of (Sv4) we have

$$E(S) = S^M(S, S_N) \leq S^M(S, T_N) \leq S^M(T, T_N) = E(T)$$

for $i = 1, ..., n$.
In similar way and by (Sv1) we may prove case $S \succeq T \succeq \mathbf{E}$.
What finished the proof (E3) and Proposition 5.

□

Example 5. If we use similarity measures from Example 3 with $\mathrm{Prec}_{\mathcal{A}_{meanL^I}}, \mathrm{Prec}_{\mathcal{A}_{meanpow}}, \mathrm{Prec}_{\mathcal{A}_{meanmax}}$, then we obtain three entropy measures satisfies the equation (12) with respect to $N(x) = [1 - \overline{x}, 1 - \underline{x}]$.

V. Application in k-NN classifiers

We use classification to evaluate the effectiveness of filling in missing data method. A classification consists of determining the class (decision) to which a new, previously undefined object should be assigned. We used k-Nearest Neighbors classifier (k-NN) [43] but we allow uncertainty in the data in an epistemic sense, i.e., the data are represented by interval values. In our approach, the object being classified belongs to the class to which most of its k-nearest neighbors belong. We present the modified algorithm IV-kNN presented in [44] of an interval-value fuzzy classifier to support decision-making processes based on imprecise (uncertain) data. The main aim was to develop a comprehensive and efficient solution to complete the uncertainty magnitude data at each step of the process.

Unlike the IV-kNN [44], the classification concept peoposed hehe allows for missing values. Therefore, it is worth pointing that our concept is innovative, in which, in the process of data completion, we proposed methods based on the entropy measure with generalized reflectivity.

A. Used methods

1) IV-kNN algorithm: The IV-kNN algorithm outlines the main steps of the proposed classification method by modifying the algorithm presented in [44] which we use in situations with missing data.

Algorithm 1: IV-kNN Algorithm

input : n element data set P composed of vectors of size a with decisions (see Section V-B); k - nearest neighbors
output: Decision class for the tested object t

▷ Step 1. Two-stage data fuzzyfication
for *all* $x \in P$ **do**
  find maximal value M_i;
  find minimal value m_i;
  for *all* $a \in x$ **do**
    $a \leftarrow \frac{a - m_i}{M_i - m_i}$
for *all* $x \in P$ **do**
  for *all* $a \in x$ **do**
    $a \leftarrow [a(1 - 0.25g(a)), a(1 - 0.25g(a)) + 0.25g(a)]$
▷ g is ignorance function
▷ Step 2. Methods refilling
for *all* $x \in P$ **do**
  fill the missing values with one of proposed method I-IV;
▷ Step 3. Take new object without decision
for *all* $a \in t$ **do**
  $a \leftarrow \frac{a - m_i}{M_i - m_i}$;
  $a \leftarrow [a(1 - 0.25g(a)), a(1 - 0.25g(a)) + 0.25g(a)]$
▷ Step 4. Calculating similarities
for *all* $x \in P$ **do**
  $sim(x) \leftarrow S(x, t)$
▷ Step 5. Sorting the sim array to select the k most similar objects
sort(sim)
▷ Step 6. Aggregation
for $i := 1, 2, \ldots, k$ **do**
  Aggregation of each classes amidst k most similar neighbors;
▷ Step 7. Class selection
Select the best interval for each class (Pekala 2020).

2) Complete missing data algorithm: **Entropy.**

If the tested object has a missing value then to complete this value we propose the following method based on entropy measure (Entropy Algorithm). For the most similar objects from the same decision class to the given object with missing value we find value of the object with the minimal entropy.

B. Structure of dataset

We used Wisconsin datasets with data of breast cancer records. This is one of the popular datasets from UCI Machine Learning Repository [46]. Data containing information on 569 instances. The data comes from a digital image of the samples taken with a fine needle aspirate (FNA) of the breast mass. The data describes the characteristics of the cell nuclei in the

Algorithm 2: Entropy Algorithm

input : Object O_z with missing value on the attribute t;
l - no. most similar objects from the same decision class;
output: Competed value $O_z(t)$;

$minEntropy = MaxInteger$; ▷ Initialization $minEntropy$ value as some big value
$noObject = -1$; ▷ storage number object with smallest entropy
for $i := 1, 2, \dots, l$ **do**
 ▷ check decision O_z is the same in O_i;
 if *dec(O_z) = dec(O_i)* **then**
  $minEntropy \leftarrow E(O_i)$;
  $noObject \leftarrow i$;
  if $\exists_{1 \leq p,q \leq l, p \neq q}$ *and* $E(O_p) = E(O_q)$ **then**
   $minEntropy \leftarrow E(O_p)$;
   $noObject \leftarrow i$;
$O_z(t) \leftarrow O_{noObject}(t)$;
 ▷ Finally, the value for the missing attribute is completed from the attribute of the won object;
if $S(O_z, O_p) > S(O_z, O_q)$ **then**
 $O_z(t) := O_p(t)$ else $O_z(t) := O_q(t)$.

image. Each records contain the parameters of the cell nucleus in the form of ten real-valued features:

- radius (mean of distances from center to points on the perimeter)
- texture (standard deviation of gray-scale values)
- perimeter
- area
- smoothness (local variation in radius lengths)
- compactness ($\frac{perimeter^2}{area-1.0}$)
- concavity (severity of concave portions of the contour)
- concave points (number of concave portions of the contour)
- symmetry
- fractal dimension ($coastline\ approximation - 1$)

A unique diagnosis is assigned to each record in the database. Diagnosis M means malignant changes, while diagnosis B means benign changes. The set contains 212 records with the type of malignant changes, and 357 objects with the benign type of lesions.

C. The conclusion of our approach

In the input processing algorithm, we focused on three aspects:

1) Adaptation of the k coefficient;
2) Appropriate selection of aggregating functions used to construct the similarity measure (including precedence measure) (Step 4 of IV-kNN Algorithm (Algorithm 1) using Proposition 3) and Example 3 and in the construction of entropy measure (Step 2, Method IV, using Propositions 4 - 5, and Examples 4 - 5) (this class of aggregations we denote by S_A). Moreover, the selection of aggregation to use in Step 6 (this class of aggregations we denote by C_A);
3) Appropriate selection of a relationship that orders the data in descending order (Step 5 and one of the cases in Step 7 of IV-kNN algorithm): $\{\leqslant_2, \leqslant_{XY}, \leqslant_{Lex1}, \leqslant_{Lex2}\}$.

We consider the effectiveness of different methods to complete missing data in our analysis in two cases:

I. We tested the following methods (see Table I):

1) **Method I** For a missing value, insert a $[0, 1]$ interval.
2) **Method II** Search for the most similar k objects and select their value for the missing attribute and calculate for the missing value insert the minimum, maximum or average of the selected values on the missing attribute: $[min(a_1, ...a_k), max(a_1, ...a_k)]$, $[avg(a_1, ...a_k), avg(a_1, ...a_k)]$.
3) **Method III** Finding the most similar l objects similarly to Method II but from the same decision class as the object with a missing value.
4) **Method IV. Entropy.** We use the Algorithm 2 (Entropy Algorithm) to fill in missing values.

II. Different entropies used in the Method IV (see Table II):

1) Proposed new entropies from Example 4 and 5 (denoted by E_I).
2) Entropy from [23]:
$$E_{II}(F) = \sum_{i=1,\dots,n} w_{F(x_i)}$$
for $F \in IVFS(X)$, and $card(X) = n$.
3) Entropy from [24]:
$E(f) = \frac{a}{b}$, where a is a distance $(f; f_{near})$ from f to the nearer point f_{near} among 1_{L^I} and 0_{L^I}, and b is the distance $(f; f_{far})$ from f to the farther point f_{far} among 1_{L^I} and 0_{L^I}. For n points belonging to an IVFS we have $E(F) = \frac{1}{n}\sum_{i=1}^{n} E(f_i)$, where $f_i = F(x_i)$. Fortunately enough, while applying the Hamming distances, the entropy of IVFSs is the ratio of the biggest cardinalities $(max \sum Counts)$ involving only F and F^c:
$$E_{III}(F) = \frac{1}{n}\sum_{i=1}^{n} \frac{maxCount(F(x_i) \wedge F^c(x_i))}{maxCount(F(x_i) \vee F^c(x_i))},$$
where $F \in IVFS(X)$. Thus concerning [24] in practice we calculate:
If $d(F, 1_{L^I}) > d(F, 0_{L^I})$, then
$$E_{III}(F) = \frac{1}{n}\sum_{i=1}^{n} \frac{\overline{F}(x_i)}{1 - \underline{F}(x_i)}$$
else
$$E_{III}(F) = \frac{1}{n}\sum_{i=1}^{n} \frac{1 - \underline{F}(x_i)}{\overline{F}(x_i)},$$
where for arbitrary $A, B \in IVFS(X)$ we have

$d(A,B) = \frac{1}{n}\sum_{i=1}^{n}(|\underline{A}(x_i) - \underline{B}(x_i)| + |\overline{A}(x_i) - \overline{B}(x_i)| + |w_A(x_i) - w_B(x_i)|)$.

In the process of the tests, it was assumed:

- The data was divided into a training part and a test part. The training part has 70% of all and the testing part has 30% from all objects.
- We test the algorithm using 10 times cross-validation.
- From the ten values obtained in this way for accuracy (ACC), sensitivity (SENS), specificity (SPEC), and precision (PREC) their arithmetic mean was calculated and presented as the value obtained by the algorithm with particular parameters.
- The algorithm was tested on data with 5%, 25%, 50% of missing values.

The tests performed in paper [47] showed that the methods of filling missing values by using our modified IV-kNN are characterized by a slower decrease in accuracy of the classification for the third method compare to the first and the second, with an increase in the number of missing values in the data. This is why we compare Method IV only with Method III in Table I. By MS we denote the percent of missing values in data. We observe progress in each aspect by the use of Method IV. To compare we choose the best representative results by the following parameters: for Method III: $C_A = \mathcal{A}_{meanpow}$, $S_A = \mathcal{A}_{mean}$ and the order $\leq_{lex2}$ and for Method IV: $C_A = S_A = \mathcal{A}_{MeanMax}$ and the order $\leq_{XY}$ and both with $k = 5$. With the parameters adopted in method IV and used in the IV-kNN classification algorithm, we see a comparison of effectiveness at different entropies apply in this completing method (Table II) and we notice an improvement of effectiveness in most parameters of classification by using new entropy introduced in the study using precedence measures. In particular, we observe progress on accuracy or sensitivity.

Method	ACC	SENS	SPEC	PREC	MS
Meth.III	0,8555894	**0,9531871**	0,7877616	0,8871398	5%
	0,8450197	**0,9529377**	0,7587463	0,8750599	25%
	0,8399786	**0,9511847**	0,7513039	0,8683283	50%
Meth.IV	**0,8875097**	0,9370335	**0,8024434**	**0,8914741**	5%
	0,8800044	0,9423466	**0,7698350**	**0,879212**	25%
	0,8778067	0,9372094	**0,7770936**	**0,8775432**	50%

TABLE I

TABLE FOR THE DIFFERENT METHOD OF FILLING MISSING VALUES.

Entropy	ACC	SENS	SPEC	PREC	MS
E_I	**0,8875097**	**0,9370335**	0,8024434	**0,8914741**	5%
	0,8800044	**0,9370166**	0,7698350	**0,879212**	25%
	0,8778067	**0,9370094**	0,7770936	0,8775432	50%
E_{II}	0,8870261	0,9369867	0,8011517	0,8907949	5%
	0,8742604	0,9352522	0,7523640	0,8717101	25%
	0,8683938	0,9342521	0,7482776	0,864492	50%
E_{III}	0,8864288	0,9346241	**0,8038857**	0,8913553	5%
	0,8808355	0,9314817	**0,777911**	0,872423	25%
	0,8768056	0,9165073	**0,8091392**	**0,8910009**	50%

TABLE II

TABLE FOR THE DIFFERENT ENTROPIES USED IN ALGORITHM.ENTROPY.

VI. CONCLUSIONS

This paper presents a new approach regarding the possibility of using the similarity measure and the entropy measure to study interval data. Introduced new methods of data analysis based on the analysis of the width of the intervals. In addition, the formula for the possibility of comparing interval data to data stored in the form of fuzzy sets has been outlined. The relations between the entropy measure and operations such as including, aggregation for interval data have been analyzed.

Similarity measures were used to determine the decision, using the k-NN algorithm from the given form of interval values. In case of missing data, methods of complementing them with the use of etropia have been proposed.

In subsequent studies, we intend to check the effectiveness of our algorithms for other data sets with missing values, and to compare our approaches to other algorithms.

REFERENCES

[1] H. Bustince, E. Barrenechea, M. Pagola, J. Fernández, Z. Xu, B. Bedregal, J. Montero, H. Hagras, F. Herrera and B. De Baets, "A Historical Account of Types of Fuzzy Sets and Their Relationships", *IEEE Transactions on Fuzzy Syst.*, vol. 24(1), pp. 174 - 194, 2016.

[2] U. Bodenhofer, B. De Baets, J. Fodor, "A compendium of fuzzy weak orders: Representations and constructions", *Fuzzy Sets Syst.,* vol. 158, pp. 811 - 829, 2007.

[3] K. Bosteels, E.E. Kerre, "On a reflexivity-preserving family of cardinality-based fuzzy comparison measures", *Inf. Sci.*, vol. 179, pp. 2342 - 2352, 2009.

[4] H. Bustince, E. Barrenechea, M. Pagola, "Image thresholding using restricted equivalence functions and maximizing the measures of similarity", *Fuzzy Sets Syst.*, vol. 158, pp. 496 - 516, 2007.

[5] B. De Baets, H. De Meyer, H. Naessens, "A class of rational cardinality-based similarity measures", *J. Comp. Appl. Math.*, vol. 132, pp. 51 - 69, 2001.

[6] B. De Baets, S. Janssens, H. De Meyer, "On the transitivity of a parametric family of cardinality-based similarity measures", *Internat. J. Appr. Reas.*, vol. 50, pp. 104 - 116, 2009.

[7] B. Jayaram, R. Mesiar, "I-Fuzzy equivalence relations and I-fuzzy partitions", *Inf. Sci.*, vol. 179, pp. 1278 - 1297, 2009.

[8] B. Pękala, "Preservation of properties of interval-valued fuzzy relations", in: Carvalho, J.P. et al. (eds.) IFSA/EUSFLAT 2009, pp. 1206–1210, 2009.

[9] E.S. Palmeira, B. Bedregal, H. Bustince, D. Paternain, L. De Miguel, "Application of two different methods for extending lattice-valued restricted equivalence functions used for constructing similarity measures on L -fuzzy sets", *Inf. Sci.*, vol. 441, pp. 95 - 112, 2018.

[10] B. Pękala, U. Bentkowska, H. Bustince, J. Fernandez, M. Galar, "Operators on intuitionistic fuzzy relations", IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), pp. 1-8, 2015.

[11] B. Pękala, "Operations on Interval Matrices", in: M. Kryszkiewicz et al. (eds.) Rough Sets and Intelligent Systems Paradigms. Springer-Verlag, Berlin, *LNAI*, vol. 4585, pp. 613-620, 2007.

[12] U. Bentkowska, J.G. Bazan, M. Mrukowicz, L. Zareba, P. Molenda, "Multi-class classification problems for the k-NN algorithm in the case of missing values", 2020 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), pp. 1-8, 2020.

[13] K. Dyczkowski, A. Wójtowicz, P. Żywica, A. Stachowiak, R. Moszyński, and S. Szubert, "An intelligent system for computer-aided ovarian tumor diagnosis', in Intelligent Systems'2014, D. Filev et al. eds., Springer, *Advances in Intelligent Systems and Computing*, vol. 323, pp. 335–343, 2015.

[14] M. Štepnicka, N. Cao, L. Behounek, M. Burda, and A. Dolný, "Missing values and dragonfly operations in fuzzy relational compositions", *Int. J. Approx. Reason.*, vol. 113, pp. 149–170, 2019.

[15] A. Wójtowicz, P. Żywica, A. Stachowiak, and K. Dyczkowski, "Solving the problem of incomplete data in medical diagnosis via interval modeling", *Applied Soft Computing*, vol. 47, pp. 424–437, 2016.

[16] P. Żywica, K. Dyczkowski, A. Wójtowicz, A. Stachowiak, S. Szubert, and R. Moszyński, "Development of a fuzzy-driven system for ovarian tumor diagnosis", *Biocybernetics and Biomedical Engineering*, vol. 36, no. 4, pp. 632–643, 2016.

[17] R. Sambuc, "Fonctions ϕ-floues: Application á l'aide au diagnostic en pathologie thyroidienne", Ph.D. Thesis, Universit*é* de Marseille, France (in French), 1975.

[18] Z. Takáč, H. Bustince, J. M. Pintor, C. Marco-Detchart and I. Couso, "Width-Based Interval-Valued Distances and Fuzzy Entropies," *IEEE Access*, vol. 7, 14044-14057, 2019, doi: 10.1109/ACCESS.2019.2893800.

[19] M.J. Asiain, H. Bustince, R. Mesiar, A. Kolesarova, Z. Takac, "Negations with respect to admissible orders in the interval-valued fuz.zy set theory", *IEEE Transactions on Fuzzy Systems*. DOI: 10.1109/TFUZZ.2017.2686372

[20] U. Bentkowska, H. Bustince, A. Jurio, M. Pagola, and B. Pękala, "Decision making with an interval-valued fuzzy preference relation and admissible orders", *Appl. Soft Comput.*, vol. 35, 792 -801, 2015.

[21] H. Bustince, "Indicator of inclusion grade for interval-valued fuzzy sets. Application to approximate reasoning based on interval-valued fuzzy sets", *Internat. J. Appr. Reas.*, vol. 23(3), pp. 137 - 209, 2000.

[22] B. Pękala, "Uncertainty Data in Interval-Valued Fuzzy Set Theory. Properties, algirithms and applications", Studies in Fuzziness and Soft Computing, Springer, 2019.

[23] P. Burillo and H. Bustince, "Entropy on intuitionistic fuzzy sets and on interval-valued fuzzy sets", *Fuzzy Sets Syst.*, vol. 78 (3), 305-316, 1996.

[24] E. Szmidt and J. Kacprzyk, "Entropy for intuitionistic fuzzy sets", *Fuzzy Sets Syst.*, vol. 118 (3), 467-477, 2001.

[25] Z. Takac, M. Minárová, J. Montero, E. Barrenechea, J. Fernandez, and H. Bustince, "Interval-valued fuzzy strong S-subsethood measures, interval-entropy and P-interval-entropy", *Inf. Sci.*, vol. 432, 97-115, 2018.

[26] L.A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning", *Inf. Sci.*, vol. 8, Part I, pp. 199 - 251, Part II, pp. 301 - 357, *Inf. Sci.*, vol. 9, Part III, pp. 43 - 80, 1975.

[27] I.B. Turksen, "Interval-valued fuzzy sets based on normal forms", *Fuzzy Sets Syst.*, vol. 20, pp. 191–210, 1986.

[28] M.B. Gorzalczany, "A method of inference in approximate reasoning based on interval-valued fuzzy sets", *Fuzzy Sets Syst.*, vol. 21, pp. 1-17, 1987.

[29] H. Bustince, J. Fernandez, A. Kolesárová, R. Mesiar, "Generation of linear orders for intervals by means of aggregation functions", *Fuzzy Sets Syst.*, vol. 220, pp. 69 - 77, 2013.

[30] H. Zapata, H. Bustince, S. Montes, B. Bedregal, G.P. Dimuro, Z. Takáč, M. Baczyński, J. Fernandez, "Interval-valued implications and interval-valued strong equality index with admissible orders", *Internat. J. Appr. Reas.*, vol. 88, pp. 91 - 109, 2017.

[31] Z.S. Xu, R.R.Yager, "Some geometric aggregation operators based on intuitionistic fuzzy sets", *Int. J. Gen. Syst.*, vol. 35, pp. 417-433, 2006.

[32] G. Beliakov, H. Bustince, T. Calvo, "A practical Guide to Averaging Functions", In: Studies in Fuzziness and Soft Computing. Springer Vol. 329, 2016.

[33] M. Komorníková, R. Mesiar, "Aggregation functions on bounded partially ordered sets and their classification", *Fuzzy Sets Syst.*, vol. 175, pp. 48 - 56, 2011.

[34] G. Deschrijver, E.E. Kerre, "Implicators based on binary aggregation operators in interval-valued fuzzy set theory", *Fuzzy Sets Syst.*, vol. 153(2), pp. 229 - 248, 2005.

[35] G. Deschrijver, "Quasi-arithmetic means and OWA functions in interval-valued and Atanassovs intuitionistic fuzzy set theory", in: EUSFLAT-LFA 2011, Aix-les-Bains, France, pp. 506 - 513, 2011.

[36] M. Elkano, J.A. Sanz, M. Galar, B. Pękala, U. Bentkowska, H. Bustince, "Composition of interval-valued fuzzy relations using aggregation functions", *Inf. Sci.*, vol. 369, pp. 690 - 703, 2016.

[37] W.Y. Zeng, P. Guo, "Normalized distance, similarity measure, inclusion measure and entropy of interval-valued fuzzy sets and their relationship", *Inf. Sci.*, vol. 178, pp. 1334 - 1342, 2008.

[38] Z. Takáč, M. Minárová, J. Montero, E. Barrenechea, J. Fernandez, H. Bustince, "Interval-valued fuzzy strong S-subsethood measures, interval-entropy and P-interval-entropy", *Inf. Sci.*, vol. 432, pp. 97 - 115, 2018.

[39] B. Pękala, K. Dyczkowski, P. Grzegorzewski, U. Bentkowska, "Inclusion and similarity measures for interval-valued fuzzy sets based on aggregation and uncertainty assessment", *Inform. Sci.*, vol. 547, 1182-1200, 2021.

[40] H. Bustince, C. Marco-Detchart, J. Fernandez, C. Wagner, J.M. Garibaldi, Z. Takáč, "Similarity between interval-valued fuzzy sets taking into account the width of the intervals and admissible orders", *Fuzzy ets and Systems*, vol. 390, pp. 23-47, 2020.

[41] B. Pękala, U. Bentkowska, M. Sesma-Sara, J. Fernandez, J. Lafuente, A. Altalhi, M. Knap, H. Bustince, "Interval subsethood measures with respect to uncertainty for interval-valued fuzzy setting", *Int. J. Comp. Int. Syst.*, vol. 3, pp. 167–177, 2020.

[42] G. Deschrijver, C. Cornelis, "Represantability in interval-valued fuzzy set theory", *International Journal of Uncertainty. Fuzziness Knowledge-Based Systems,* vol. 15(3), pp. 345–361, 2007.

[43] T. M. Cover and P. E. Hart, "Nearest neighbor pattern classification", *IEEE Transactions on Information Theory*, vol. 13(1), 21–27, 1967.

[44] B. Pękala, E. Rak, D. Kosior, M. Mrukowicz, J.G. Bazan, "Application of similarity measures with uncertainty in classification methods", 2020 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), pp. 1-8, 2020.

[45] E. Barrenechea, J. Fernandez, M. Pagola, F. Chiclana, H. Bustince, "Construction of interval-valued fuzzy preference relations from ignorance functions and fuzzy preference relations: Application to decision making", Knowledge-Based Systems, 58, pp. 33–44, 2014.

[46] https://archive.ics.uci.edu/ml/index.php

[47] D. Kosior, B. Pękala, "Influence of interval-valued measures on classification methods with missing values", IWIFSGN'2020, in review.