Application of similarity measures with uncertainty in classification methods

1st Barbara Pękala, University of Rzeszów, Rzeszów, Poland, bpekala@ur.edu.pl

2nd Ewa Rak, University of Rzeszów, Rzeszów, Poland, ewarak@ur.edu.pl

3rd Dawid Kosior, University of Rzeszów, Rzeszów, Poland, kosiordawid@gmail.com

4th Marcin Mrukowicz, University of Rzeszów, Rzeszów, Poland, mmrukowicz@ur.edu.pl

5th Jan G. Bazan, University of Rzeszów, Rzeszów, Poland, bazan@ur.edu.pl

***Abstract*—In this paper, the problem of measuring the degree of inclusion and similarity measure for interval-valued fuzzy sets is considered. We recall inclusion and similarity measures with uncertainty by using the partial or linear order on interval-valued fuzzy sets. Moreover, we discuss an influence of inclusion and similarity measures with uncertainty to decision-making algorithm that uses those new measures.**

I. Introduction

Many new approaches and theories for investigating and modeling imprecision and uncertainty have been proposed since fuzzy sets were introduced by Zadeh [1]. Interval-valued fuzzy sets (Zadeh, Sambuc 1975) like intuitionistic fuzzy sets (Atanassov 1986) appeared very useful because of their flexibility (for more details see [2]). Moreover, various applications of interval-valued fuzzy sets for solving real-life problems involving pattern recognition, medical diagnosis and decision-making or image thresholding were successfully proposed. The above mentioned results have been enabled due to the substantial progress in the theory of interval-valued fuzzy sets. For example, many researchers proposed various distances, measures of inclusion, measures of equivalence or similarity measures for interval-valued fuzzy sets and examined different types of relations between them ([3]-[9]).

The main motivation of the present paper is to use the degree of inclusion and similarity between interval-valued fuzzy sets to classification problem, especially to the k-nearest neighbors (k-NN) algorithm. A fundamental novelty of the suggested approach is to make use of the ordering between intervals and the width of those intervals representing both uncertainty of information and imprecision of the membership functions. In addition, the use of different types of aggregation functions was also helpful here.

The paper is organized as follows. In Section 2 basic information on interval-valued fuzzy sets are recalled. Inclusion and similarity degree measures for interval-valued fuzzy sets are presented in Section 3. Finally, the algorithm of application new similarity measures in k-NN classifiers is considered (Section 4).

II. Interval-valued fuzzy setting

Let $L^I = \{[\underline{a}, \overline{a}] :\ \underline{a}, \overline{a} \in [0,1],\ \underline{a} \leqslant \overline{a}\}$ denote a family of all compact subintervals of the unit interval. Let $X \neq \emptyset$ denote a universe of discourse. According to the following papers by Zadeh [10], Sambuc [11], Turksen [12] and Gorzalczany [13] we define an **interval-valued fuzzy set** A in X as a mapping $A : X \to L^I$ such that for each $x \in X$

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

attributes the degree of membership of an element x into A. Furthermore, the family of all interval-valued fuzzy sets in X will be denoted by $\mathrm{IVFS}(X)$. We will assume hereinafter that the considered universe of discourse is finite, i.e. $X = \{x_1, \dots, x_n\}$. Indeed, for fuzzy set the membership of each element x is always a precisely given real number. On the other hand, in the case of the interval-valued fuzzy set the membership of an element x is not precisely done. Here we can only indicate an upper and lower bound of its possible membership. And this is why the interval-valued fuzzy sets appear so useful for approximate reasoning. Obviously, each fuzzy set A could be treated as the interval-valued fuzzy set such that $\underline{A}(x) = \overline{A}(x)\ \forall x \in X$. Thus, $\mathrm{P}^X \subset \mathrm{FS}(X) \subset \mathrm{IVFS}(X)$, where $\mathrm{FS}(X)$ stands for a family of all fuzzy sets in X while P^X is the class of all crisp subsets of X.

We define basic operations (intersection, union and complement) as follows for $x \in X$:

$$A \cap B = \big\{\langle x, \big[\min\{\underline{A}(x), \underline{B}(x)\}, \min\{\overline{A}(x), \overline{B}(x)\}\big]\rangle\big\}, \quad (1)$$

$$A \cup B = \big\{\langle x, \big[\max\{\underline{A}(x), \underline{B}(x)\}, \max\{\overline{A}(x), \overline{B}(x)\}\big]\rangle\big\}, \quad (2)$$

$$A^c = \big\{\langle x, \big[1 - \overline{A}(x), 1 - \underline{A}(x)\big]\rangle\big\}. \quad (3)$$

Obviously, instead of $\min$, $\max$ and the standard negation one may use any t-norm, s-conorm and other fuzzy negation in (1)–(3). One can prove that $(\mathrm{IVFS}(X), \cap, \cup)$ is a distributive lattice which satisfies the Morgan's laws.

Before introducing inclusion and similarity indicators we have to consider orderings and aggregation operators in L^I.

A. Orders in the interval setting

Consider two interval-valued fuzzy sets A and B in the same universe of discourse X. Since in this section we restrict our attention to possible relations between $A(x)$ and $B(x)$ for any fixed $x \in X$ let us adopt the following simplified notation $A(x) = [\underline{A}(x), \overline{A}(x)] = [\underline{a}, \overline{a}]$ and $B(x) = [\underline{B}(x), \overline{B}(x)] = [\underline{b}, \overline{b}]$.

The best known partial order in L^I is defined as follows

$$[\underline{a}, \overline{a}] \leqslant_2 [\underline{b}, \overline{b}] \Leftrightarrow \underline{a} \leqslant \underline{b} \quad \text{and} \quad \overline{a} \leqslant \overline{b}, \quad (4)$$

978-1-7281-6932-3/20/$31.00 ©2020 IEEE

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

The operations joint and meet are defined in L^I as follows

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

The structure $(L^I, \vee, \wedge)$ is a complete lattice, with the partial order $\leqslant_2$. Obviously, $1_{L^I} = [1,1]$ and $0_{L^I} = [0,0]$ are the greatest and the smallest element of $(L^I, \leqslant_2)$, respectively.

Since in many real-life problems we need a linear order to be able to compare any two intervals, we are interested in extending the partial order $\leqslant_2$ to a linear one. The concept of, so called, admissible order would be of useful there.

Definition 1 ([14], Def. 3.1)**.** An order $\leq_{Adm}$ in L^I is called **admissible** if $\leq_{Adm}$ is linear in L^I and for all $a, b \in L^I$ $a \leq_{Adm} b$ whenever $a \leqslant_2 b$.

Admissible orders were further studied, e.g. in [15] or [16]. A construction of admissible linear orders based on aggregation functions was given in [14].

Proposition 1 ([14], Prop. 3.2)**.** *Let $\Psi, \Upsilon : [0,1]^2 \to [0,1]$ be two continuous aggregation functions, such that, for all $a = [\underline{a},\overline{a}], b = [\underline{b},\overline{b}] \in L^I$, the equalities $\Psi(\underline{a},\overline{a}) = \Psi(\underline{b},\overline{b})$ and $\Upsilon(\underline{a},\overline{a}) = \Upsilon(\underline{b},\overline{b})$ hold if and only if $a = b$. If the order $\leqslant_{\Psi,\Upsilon}$ on L^I is defined by*

$$a \leqslant_{\Psi,\Upsilon} b \Leftrightarrow \Psi(\underline{a},\overline{a}) < \Psi(\underline{b},\overline{b})$$
$$\textit{or } (\Psi(\underline{a},\overline{a}) = \Psi(\underline{b},\overline{b}) \textit{ and } \Upsilon(\underline{a},\overline{a}) \leqslant \Upsilon(\underline{b},\overline{b})), \quad (5)$$

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

Further on the notation $<_{\Psi,\Upsilon}$ would indicate that in the strict inequality holds in (5).

Here are some natural examples of admissible orders in L^I (see, e.g., [14]):

- the Xu-Yager order [17]

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

- lexicographical orders

$$[\underline{a},\overline{a}] \leqslant_{Lex1} [\underline{b},\overline{b}] \Leftrightarrow \underline{a} < \underline{b} \text{ or } (\underline{a} = \underline{b} \text{ and } \overline{a} \leqslant \overline{b}) \quad (6)$$
$$[\underline{a},\overline{a}] \leqslant_{Lex2} [\underline{b},\overline{b}] \Leftrightarrow \overline{a} < \overline{b} \text{ or } (\overline{a} = \overline{b} \text{ and } \underline{a} \leqslant \underline{b}) \quad (7)$$

- the $\alpha\beta$ order

$$a \leqslant_{\alpha\beta} b \Leftrightarrow K_\alpha(\underline{a},\overline{a}) < K_\alpha(\underline{b},\overline{b}) \quad (8)$$
$$\text{or } (K_\alpha(\underline{a},\overline{a}) = K_\alpha(\underline{b},\overline{b}) \text{ and } K_\beta(\underline{a},\overline{a}) \leqslant K_\beta(\underline{b},\overline{b})),$$

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

It is worth noting that the orders $\leqslant_{XY}$, $\leqslant_{Lex1}$ and $\leqslant_{Lex2}$ are special cases of $\leqslant_{\alpha\beta}$ with $\leqslant_{0.5\beta}$ (for $\beta > 0.5$), $\leqslant_{1,0}$, $\leqslant_{0,1}$, respectively.

Moreover, $\leqslant_{XY}$, $\leqslant_{Lex1}$, $\leqslant_{Lex2}$ and $\leqslant_{\alpha\beta}$ are $\leqslant_{\Psi,\Upsilon}$ orders defined by pairs of weighted means (cf. Proposition 1). In the case of $\leqslant_{Lex1}$ and $\leqslant_{Lex2}$ these means are reduced to the pairs of projections: P_1, P_2 and P_2, P_1, respectively.

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

B. Interval-valued aggregation functions

Now we recall the concept of an aggregation function on L^I which is an important notion in many applications. We consider aggregation functions both with respect to $\leqslant_2$ and $\leqslant_{Adm}$.

Remark 2. In the later part of the paper we will use the notation $\leq$ both for the partial or admissible linear order, with 0_{L^I} and 1_{L^I} as minimal and maximal element of L^I, respectively. Regarding the results for the partial order, the previously introduced notation $\leq_2$ will be used while the results for the admissible linear orders will be used with the notation $\leq_{Adm}$ (sometimes with the appropriate example of such admissible linear order).

Definition 2 ([16], [18], [19])**.** Let $n \in \mathbf{N}$, $n \geqslant 2$. An operation $\mathcal{A} : (L^I)^n \to L^I$ is called an interval-valued aggregation function if it is increasing with respect to the order $\leq$ (partial or linear (see Remark 2)), i.e.

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

The special case of interval-valued aggregation operation is a representable interval-valued aggregation function with respect to $\leq_2$.

Definition 3 ([20], [21])**.** An interval-valued aggregation function $\mathcal{A} : (L^I)^n \to L^I$ is called representable if there exist aggregation functions $A_1, A_2 : [0,1]^n \to [0,1]$ such that

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

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

The next result shows the characterization of representable aggregation functions on L^I.

Theorem 1 ([22])**.** *An operation $\mathcal{A} : (L^I)^n \to L^I$ is a representable interval-valued aggregation function with respect to $\leq_2$ if and only if there exist aggregation functions $A_1, A_2 : [0,1]^n \to [0,1]$ such that for all $x_1, ..., x_n \in L^I$ and $A_1 \leqslant A_2$*

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

Example 1. Lattice operations $\wedge$ and $\vee$ on L^I are representable aggregation functions on L^I with $A_1 = A_2 = \min$ in the first case and $A_1 = A_2 = \max$ in the second one. It holds true with respect to the order $\leq_2$, but not $\leq_{Lex1}$, $\leq_{Lex2}$ or $\leq_{XY}$. Moreover, many other examples of representable aggregation functions with respect to $\leq_2$ may be considered, such as:

- 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
  $\mathcal{A}_{gmean}([\underline{x},\overline{x}],[\underline{y},\overline{y}])=[\sqrt{\underline{x}\underline{y}},\sqrt{\overline{x}\overline{y}}]$,
- 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
  $\mathcal{A}_{prod}([\underline{x},\overline{x}],[\underline{y},\overline{y}])=[\underline{x}\underline{y},\overline{x}\overline{y}]$,
- the representable prod-mean
  $\mathcal{A}_{prodmean}([\underline{x},\overline{x}],[\underline{y},\overline{y}])=[\underline{x}\underline{y},\frac{\overline{x}+\overline{y}}{2}]$,
- the representable mean-max
  $\mathcal{A}_{meanmax}([\underline{x},\overline{x}],[\underline{y},\overline{y}])=[\frac{\underline{x}+\underline{y}}{2},\max\overline{x},\overline{y}]$

for $[\underline{x},\overline{x}],[\underline{y},\overline{y}]\in L^I$.

Representability is not the only possible way to build interval-valued aggregation functions with respect to $\leq_2$ or $\leq_{Adm}$. Note that $\mathcal{A}_{mean}$ is the aggregation with respect to $\leqslant_{Lex1}$, $\leqslant_{Lex2}$ and $\leqslant_{XY}$, not only to $\leqslant_2$. Moreover,

Example 2. Let $A:[0,1]^2\to[0,1]$ be an aggregation function.

- The function $\mathcal{A}_2:(L^I)^2\to L^I$ ([23]), where

$$\mathcal{A}_2(x,y)=\begin{cases}[1,1], & if\ (x,y)=([1,1],[1,1])\\ [0,A(\underline{x},\underline{y})], & otherwise\end{cases}$$

is the interval-valued aggregation function (non-representable) with respect to $\leq_{Lex1}$.

- The function $\mathcal{A}_3:(L^I)^2\to L^I$ ([23]), where

$$\mathcal{A}_3(x,y)=\begin{cases}[0,0], & if\ (x,y)=([0,0],[0,0])\\ [A(\overline{x},\overline{y}),1], & otherwise\end{cases}$$

is the interval-valued aggregation function (non-representable) with respect to $\leq_{Lex2}$.

III. Inclusion and similarity degree measures for interval-valued fuzzy sets

The inclusion measures, also known as subsethood measures, have been studied mainly by constructive approaches and axiomatic approaches. The inclusion measure has also been introduced successfully into fuzzy sets theory and their extensions. For fuzzy sets A and B a measure of fuzzy set inclusion of A in B is defined as a subset of A in B. Many researchers tried to relax the rigidity of Zadeh definition of inclusion to get a soft approach which is more compatible with the spirit of fuzzy logic. Zhang and Leung (1996) thought that quantitative methods were the main approaches in uncertainty inference which is a key problem for artificial intelligence. Thus, they presented a generalized definition for the inclusion measure, called *including degree*, to represent and measure the uncertainty information. Instead of binary discrimination, being or not being a subset [16], [24], [25], [26], several indicators giving the degree to which an interval-valued fuzzy set is a subset of another interval-valued fuzzy set were proposed. For each $A,B\in IVFS(X)$, we will represent the inclusion grade indicator of set A in set B by the measure of inclusion between their elements, i.e. intervals. This fact has led us to establish the next considerations of inclusion measure in interval setting.

A. Precedence indicator

Before giving a new perspective on measuring inclusion or similarity in the interval-valued fuzzy set environment, we have to introduce another tool useful for handling of intervals. We propose the following notion of an inclusion measure using linear or partial order and uncertainty measure/width of intervals, i.e. $w(a)=\overline{a}-\underline{a}$ denote the width of $a\in L^I$.

We consider the notion of a precedence indicator where strong inequality between inputs results in the same values of the inclusion measure for these inputs.

Moreover, we recall that $<$ and $\leq$ satisfy Remark 1 and 2.

Definition 4 ([27])**.** A function $\mathrm{Prec}:(L^I)^2\to L^I$ is said to be a **precedence indicator** if it satisfies the following conditions for any $a,b,c\in L^I$:

P1 if $a=1_{L^I}$ and $b=0_{L^I}$ then $\mathrm{Prec}(a,b)=0_{L^I}$;
P2 if $a<b$, then $\mathrm{Prec}(a,b)=1_{L^I}$ for any $a,b\in L^I$;
P3 $\mathrm{Prec}(a,a)=[1-w(a),1]$ for any $a\in L^I$;
P4 if $a\leq b\leq c$ and $w(a)=w(b)=w(c)$, then $\mathrm{Prec}(c,a)\leq\mathrm{Prec}(b,a)$ and $\mathrm{Prec}(c,a)\leq\mathrm{Prec}(c,b)$, for any $a,b,c\in L^I$.

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

Here are examples of the constructions of precedence indicator fulfilling Definition 4.

Proposition 2 ([27])**.** *For $a,b\in L^I$ the operation $\mathrm{Prec}_{\mathcal{A}}:(L^I)^2\to L^I$ is the precedence indicator*

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

for $a,b\in L^I$ and interval-valued 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. [28], [29]), such that $N_{IV}(a)=[n(\overline{a}),n(\underline{a})]\leq[1-\overline{a},1-\underline{a}]$, where n is a fuzzy negation and $\mathcal{A}$ is a representable interval-valued aggregation such that $\mathcal{A}\leq\vee$.

Similarly to the method of constructing the precedence indicator presented in [30] we get

Proposition 3. *The operation*

$$\mathrm{Prec}_w(a,b)=\begin{cases}1_{L^I}, & a<b,\\ [1-\max(w(a),r(a,b)),1-r(a,b)], & else\end{cases}$$

is the precedence indicator with respect to $\leq$, where $r(a,b)=\max\{|\underline{a}-\underline{b}|,|\overline{a}-\overline{b}|\}$ for $a,b\in L^I$.

B. Similarity measure

In this part we study a class of similarity measures between interval-valued fuzzy sets. The inspiration of this approach is firstly the fact that we develop all the notions with respect to total orders of intervals, and secondly, that we take into account the width of the intervals in such a way that the uncertainty of the output is strongly related to the uncertainty of the input. To construct the new interval-valued similarity, interval-valued aggregation functions and interval-valued inclusion measure which take into account the width of the intervals are needed.

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

$$A \preceq B \Leftrightarrow a_i \leq b_i$$

for $i = 1, ..., n$, where $\leq$ is the same kind of orders (partial or linear, see Remark 1 and 2) for each i and $a_i = A(x_i)$, $b_i = B(x_i)$. Let us note that if for $i = 1, ..., n$ we consider the same linear order $a_i \leq b_i$, then the order $A \preceq B$ between interval-valued fuzzy sets A, B is the partial one but it need not be the linear one.

We consider the following notion of strict order between interval-valued fuzzy sets $A \prec B \Leftrightarrow a_i < b_i$ for $i = 1, ..., n$ and we denote: $w(A) = (w(a_1), ..., w(a_n))$.

Definition 5 ([31]). Let $A_1 : [0,1]^n \to [0,1]$ be an aggregation function. Then operation $S : IVFS(X) \times IVFS(X) \to L^I$, which satisfies the following items:
(SIMV1) $S(A, B) = S(B, A)$ for $A, B \in IVFS(X)$;
(SIMV2) $S(A, A) = [1 - A_1(w_A(x_1), ..., w_A(x_n)), 1]$;
(SIMV3) $S(A, B) = 0_{L^I}$, if $\{A(x_i), B(x_i)\} = \{0_{L^I}, 1_{L^I}\}$;
(SIMV4) if $A \preceq B \preceq C$ and $w_A(x_i) = w_B(x_i) = w_C(x_i)$, then $S(A, C) \leq S(A, B)$ and $S(A, C) \leq S(B, C)$
is called a similarity measure for $i = 1, ..., n$..

Proposition 4. *Let* Prec *be an inclusion measure. If* $\mathcal{A} = [A_1, A_2]$, $\mathcal{B} = [B_1, B_2]$ *are representable interval-valued aggregation functions for which* A_1 *is as in Definition 5 and self-dual,* $\mathcal{B}$ *is symmetric with the neutral element* 1_{L^I} *and* B_1 *is idempotent aggregation function, then the operation* $S : IVFS(X) \times IVFS(X) \to L^I$*:*

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

is a similarity measure, where $w_A(x_i) = w_B(x_i)$.

Proof. Let $A, B, C \in IVFS(X)$. Symmetry of S holds by symmetry of $\mathcal{B}$, so (SIMV1) of Definition 5 is fulfilled. For $A = B$ and by self duality of A_1 and idempotency of B_1, we have
$S(A, A) = \mathcal{A}_{i=1}^n(\mathcal{B}(\text{Prec}(A(x_i), A(x_i)), \text{Prec}(A(x_i), A(x_i))))$
$= \mathcal{A}_{i=1}^n(\mathcal{B}([1 - w_A(x_i), 1], [1 - w_A(x_i), 1]))$
$= \mathcal{A}_{i=1}^n([B_1(1 - w_A(x_i), 1 - w_A(x_i)), B_2(1, 1)])$
$= \mathcal{A}_{i=1}^n([1 - w_A(x_i), 1]) = [1 - A_1(w_A(x_1), ..., w_A(x_n)), 1]$,
which proves (SIMV2) of Definition 5.
If ($A = 1_{L^I}$ and $B = 0_{L^I}$) or ($A = 0_{L^I}$ and $B = 1_{L^I}$), then by Definition 4 we have
$\text{Prec}(A(x_i), B(x_i)) = 0_{L^I}$ or $\text{Prec}(B(x_i), A(x_i)) = 0_{L^I}$. Thus

$$\mathcal{A}_{i=1}^n(\mathcal{B}(\text{Prec}(A(x_i), B(x_i)), \text{Prec}(B(x_i), A(x_i)))) = 0_{L^I},$$

because $\mathcal{B}$ has the neutral element 1_{L^I}, so as consequence the zero element 0_{L^I}.
In the proof of (SIMV4) from Definition 5 we consider the following cases by properties of Inc:
4.1. If $A \prec B \prec C$, then
$S(A, C) = \mathcal{A}_{i=1}^n(\mathcal{B}(\text{Prec}(A(x_i), C(x_i)), \text{Prec}(C(x_i), A(x_i))))$
$= \mathcal{A}_{i=1}^n(\mathcal{B}(1_{L^I}, \text{Prec}(C(x_i), A(x_i))))$
$= \mathcal{A}_{i=1}^n(\mathcal{B}(\text{Prec}(A(x_i), B(x_i)), \text{Prec}(C(x_i), A(x_i))))$
$\leq \mathcal{A}_{i=1}^n(\mathcal{B}(\text{Prec}(A(x_i), B(x_i)), \text{Prec}(B(x_i), A(x_i))))$
$= S(A, B)$ and $S(A, C) = S(C, A) =$
$= \mathcal{A}_{i=1}^n(\mathcal{B}(\text{Prec}(C(x_i), A(x_i)), \text{Prec}(A(x_i), C(x_i))))$
$= \mathcal{A}_{i=1}^n(\mathcal{B}(\text{Prec}(C(x_i), A(x_i)), 1_{L^I}))$
$= \mathcal{A}_{i=1}^n(\mathcal{B}(\text{Prec}(C(x_i), A(x_i)), \text{Prec}(B(x_i), C(x_i))))$
$\leq \mathcal{A}_{i=1}^n(\mathcal{B}(\text{Prec}(C(x_i), B(x_i)), \text{Prec}(B(x_i), C(x_i))))$
$= S(B, C)$;
4.2. If $A = B \prec C$, then due to the existence of the neutral element 1_{L^I} for $\mathcal{B}$ and self-duality of A_1 we have
$S(A, C) = \mathcal{A}_{i=1}^n(\mathcal{B}(\text{Prec}(A(x_i), C(x_i)), \text{Prec}(C(x_i), A(x_i))))$
$= \mathcal{A}_{i=1}^n(\mathcal{B}(1_{L^I}, \text{Prec}(C(x_i), A(x_i))))$
$\leq \mathcal{A}_{i=1}^n(\mathcal{B}(1_{L^I}, \text{Prec}(A(x_i), A(x_i))))$
$= \mathcal{A}_{i=1}^n(\text{Prec}(A(x_i), A(x_i))) =$
$[1 - A_1(w_A(x_1), ..., w_A(x_n)), 1] = S(A, B)$ and
$S(B, C) = S(A, C)$;
Similarly we may prove the case, where $A \prec B = C$ and for $A = B = C$ condition (SIMV4) of Definition 5 is obvious, which ends the proof. ☐

Directly from the above proposition we obtain

Corollary 1. *Let* Prec *be an inclusion measure. If* $\mathcal{A} = [A_1, A_2], \mathcal{B} = [B_1, B_2]$ *are representable interval-valued aggregation functions and* ${A_1}^d$ *is used in Definition 5 (SIMV2) and* $\mathcal{B}$ *is symmetric with the neutral element* 1_{L^I} *and* B_1 *is idempotent aggregation function, then the operation* $S : IVFS(X) \times IVFS(X) \to L^I$*:*

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

is a similarity measure, where $w_A(x_i) = w_B(x_i)$.

Corollary 2. *Let* Prec *be an inclusion measure. If* $\mathcal{A}$ *is symmetric representable interval-valued aggregation function with neutral element* 1_{L^I} *and* A_1 *is self-dual and idempotent, then the operation* $S : IVFS(X) \times IVFS(X) \to L^I$*:*

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

is a similarity measure.

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

$$S(A, B) = \mathcal{A}_{i=1}^n(\text{Prec}_w(A(x_i), B(x_i) \wedge \text{Prec}_w(B(x_i), A(x_i))))$$

is a similarity measure, where
$\mathcal{A} \in \{\mathcal{A}_{mean}, \mathcal{A}_{meanpow}, \mathcal{A}_{meanmax}\}$.

IV. Application in k-NN classifiers

The classification problem consists in determining the class (category) to which a new, previously unknown object should be assigned. The classifier is constructed using a training set containing data about objects for which their belonging to the class is known. These objects are described using various attributes. To assess the effectiveness of a classifier, it is used test set containing instances not known when it was created. Various classification methods have been used, among others in such important areas as image processing or medical diagnosis. The classification problem becomes significantly more complicated if we allow incomplete or uncertainty in the data. Here we consider uncertainty in epistemic sense, i.e. data are represents by interval values. In such conditions, the design of an effective classifier using classic methods can be very difficult or even impossible. Among the most commonly used techniques for classification can be distinguished k-Nearest Neighbors classifier (k-NN) [32]. It is one of the most popular non-parametric supervised learning methods being also one of the top ten algorithms in Data Mining [33] as an integral part of many applications of Machine Learning in various domains [34]. In this strategy, the object subjected to the classification process belongs to the class to which most of its k nearest neighbors belong. The nearest neighbors of the classified object should be understood as the objects from the reference set that are the most similar to it in terms of the adopted similarity measure. This principle can be formulated briefly: ”you are just like your surroundings.” Classification takes place directly, based on the vote of the object classes most similar to the classified object. This procedure guarantees that all dimensions of the data space in which the similarity is calculated are of equal importance. In this paper we present the novel concept of an interval-valued fuzzy classifier for supporting decision-making processes based on imprecise (uncertain) data. The main goal was to develop a comprehensive and effective approach that enables the modeling and processing of input data, and then the presentation of results, to be done in a way that preserves the valuable information concerning the amount of uncertainty at each stage of the process. In contrast to the fuzzy k-NN (see e.g. [35]), the proposed here concept of classification is based on the new definition of a similarity measure, related to the width of the intervals. Moreover, one of the classification problems concerns the ranking of intervals that is not usually clearly defined. We solved this problem. In turn, the problem of choosing the value of the parameter k is analogous to that in the case of the classical method of the k-nearest neighbors. We aggregate individual classes among k most similar neighbors and then select the class for the test object. Class selection is done by selecting the largest interval for each class by use the new method. The feature that distinguishes the proposed classification method is complete support for data uncertainty, i.e. we create intervals by use the ignorance functions. The use of the new similarity measure in this classification method has eliminated the problem of epistemic uncertainty, both in the learning set and during the classification. In addition, the result of the obtained interval classifier is both an indication of the class to which the new uncertain object should be assigned and a interval-valued fuzzy set describing its belonging to all known classes. It is worth emphasizing, therefore, that our approach is innovative which we proposed compared to methods based on a distance measure or a similarity measure without the generalized reflexivity condition (see e.g. [36], [37]).

A. Proposed method

The diagram IV-kNN shows the main steps of the proposed classification method.

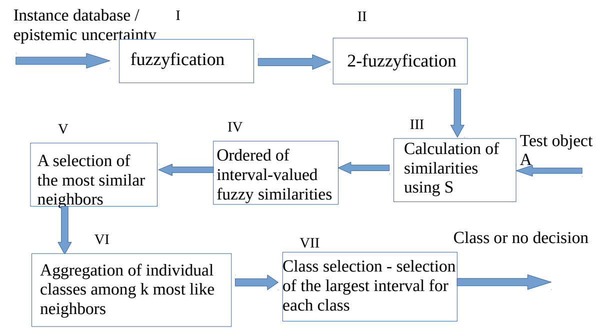

Fig. 1. Algorithm IV-kNN

I. The real data are normalized to $[0, 1]$ by classical formula based on minimal values m_i and maximal M_i for $i - th$ attribute with real value b_i

$$a_i = \frac{b_i - m_i}{M_i - m_i}.$$

II. Interval-valued fuzzy set for each instance/object is built in the following way: for each values of attribute we use for each fuzzy set, we create its corresponding interval-valued fuzzy set by using the construction method, which is as follows: We consider a fuzzy set $A \in FS(U)$ and a weak ignorance function g (i.e. a continuous function $g : [0, 1] \to [0, 1]$ such that $g(0) = g(1) = 0$, $g(0.5) = 1$ and $g(x) = g(1 - x)$ for every $x \in [0, 1]$). If for each $u_i \in U$ we take $g(\mu_A(u_i))$, $\delta(u_i)$, $\gamma(u_i) \in [0, 1]$, then the set

$$A_{IV} = \{(u_i, A_{IV}(u_i))|u_i \in U\},$$

where $A_{IV}(u_i) = G(\mu_A(u_i), g(\mu_A(u_i)), \delta(u_i), \gamma(u_i))$ is an interval-valued fuzzy set on U. Here $\delta(u_i)$, $\gamma(u_i) \in [0, 1]$ and $G(x, y, \delta, \gamma) = [x \cdot (1 - \delta \cdot y), x \cdot (1 - \delta \cdot y) + \gamma \cdot y)]$. In our case, the parameters, δ and γ are set to 0.25 as suggested by the authors ($\delta(u_i) = \gamma(u_i) = 0.25$ for all $u_i \in U$).

The following function $g(x) = 2 \cdot \min(x, 1 - x)$ for all $x \in [0, 1]$ is a weak ignorance function.

In [38], the following method to build intervals from a number and the application of the ignorance function to that number may be found. Given $x \in [0, 1]$ and the function g we have:

$$[x(1 - 0.25g(x)), x(1 - 0.25g(x)) + 0.25g(x)] \in L^I.$$

Note that the length of the intervals is equal to $g(x)$.

III. In this step of the algorithm we use one of the similarity measure of Example 3 to measure of similarity tested object with each the other. The obtained interval values may have different widths, therefore they take into account the uncertainty.

IV. We make selection of interval-valued fuzzy similarities by use one of the order

$$\{\leqslant_2, \leqslant_{XY}, \leqslant_{Lex1}, \leqslant_{Lex2}\} = \mathcal{O}.$$

V. Selection of the most similar neighbors with respect to given k, $k \in \{1, ..., n\}$, $n \in N$.

VI. We aggregate values of similarity measure with element from each class separate. We use aggregation from Example 1 or 2 with respect the same order as in point IV, e.g. $\mathcal{A}_{mean}, \vee, \wedge, \mathcal{A}_{prodmean}, \mathcal{A}_{meanmax}$.

VII. For the two intervals obtained corresponding to each class we use the following method leading to the decision, where for the classes "0" and "1" intervals are denoted respectively k_0 and k_1, in addition, we may refrain from making decisions "No Decision". The method of searching for decisions is based on comparing the width of the intervals (the smaller the width is the better) and the position of the intervals relative to the value 0.5 (we assume that the greater value of the ends of the interval than 0.5, respectively, suggests better decision clarity):

If $\mathbf{w(k_1) < w(k_0)}$ then
If $\underline{k_1} \geqslant 0.5$ then the class "1" wins
else If $\underline{k_0} \geqslant 0.5$ then the class "0" wins
else If $\overline{k_1} \geqslant 0.5$ then the class "1" wins
else If $\overline{k_0} \geqslant 0.5$ then the class "0" wins
else "No Decision"

Otherwise,
If $\mathbf{w(k_1) > w(k_0)}$ then
If $\underline{k_0} \geqslant 0.5$ then the class "0" wins
else If $\underline{k_1} \geqslant 0.5$ then the class "1" wins
else If $\overline{k_0} \geqslant 0.5$ then the class "0" wins
else If $\overline{k_1} \geqslant 0.5$ then the class "1" wins
else "No Decision"

Moreover,
If $\mathbf{w(k_1) == w(k_0)}$ then
If $k_1 \leq k_0$ $\qquad \leq \in \mathcal{O}$
then If $\underline{k_0} \geqslant 0.5$ then the class "0" wins
else If $\underline{k_1} \geqslant 0.5$ then the class "1" wins
else If $\overline{k_0} \geqslant 0.5$ then the class "0" wins
else If $\overline{k_1} \geqslant 0.5$ then the class "1" wins
else "No Decision"
If $k_0 < k_1$ $\qquad \leq \in \mathcal{O}$
then If $\underline{k_1} \geqslant 0.5$ then the class "1" wins
else If $\underline{k_0} \geqslant 0.5$ then the class "0" wins
else If $\overline{k_1} \geqslant 0.5$ then the class "1" wins
else If $\overline{k_0} \geqslant 0.5$ then the class "0" wins
else "No Decision"

B. Data set description

The dataset is a wisconsin (diagnostic) breast cancer dataset. This is one of the popular datasets from UCI Machine Learning Repository [39]. Data are from November 1995. The authors are Dr William H. Wolberg, W. Nick Street, Olvi L. Mangasarian. Data containing information on 569 instances. Each of them is represented by 32 attributes. The first one is the patient identifier, it does not carry any information so it does not participate in testing our algorithm. Features are computed from a digitized image of a fine needle aspirate (FNA) of a breast mass. They describe characteristics of the cell nuclei present in the image.
Ten real-valued features are computed for each cell nucleus:

- 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 ($coastline\ approximation - 1$)

Conditional attributes are: mean radius, mean texture, mean perimeter, mean area, mean smoothness, mean compactness, mean concavity, mean concave points, mean symmetry, mean fractal dimension, radius error, texture error, perimeter error, area error, smoothness error, compactness error, concavity error, concave points error, symmetry error, fractal dimension error, worst radius, worst texture, worst perimeter, worst area, worst smoothness, worst compactness, worst concavity, worst concave points, worst symmetry, worst fractal dimension. The decision attribute stores information about the diagnosis: malignant or benign represented by values of 0 or 1. The dataset consists of 212 objects with malignant diagnosis and 357 objects with benign diagnosis.

C. Results and discussion

In the first phase of our research we will examine the complete data, in the next (future paper) with missing information/values. Therefore at the beginning, the dataset was fuzzyfied as described in I and II. After that the dataset was splitted ten times to equal train and test parts. The implementation of the algorithm was run ten times for the selected k (we considered $k \in \{1, 3, 5\}$ for better decidability) and for each combination of orders, precedence indicators and aggregations. Each time other pair of training and test data was used. Based on confusion matrix the accuracy, sensitivity, specificity and precision were computed. At the end this measures were averaged. It is worth noting that we do not taking no decision into consideration when computing classification quality measures (they are not numerous and have no real influence to overall results). The main results obtained are listed in the tables below (however we are listing only results, which are greater than 0.5 on all measures). The following shortcuts are used in the table description: k means the number of nearest neighbors, $S_{\mathcal{A}}$ means the similarity built from a given aggregation function $\mathcal{A}$ (see Example 4), O represents the order.

We will focus on presenting the conclusions of the algorithm analysis in three aspects:

1) modification of the value of k;
2) the selection of aggregation in the construction of the similarity measure (Step III of IV-kNN algorithm using Proposition 4 and Example 3);
3) the selection of ordering relations (Step IV and one of the cases in Step VII of IV-kNN algorithm).

In the Step VI we use the same aggregation function as in Step III. Moreover, if among the most similar elements there will be elements from only one class, we compare their aggregate value in the Step VII with the interval $[0, 0]$.

First, we consider $Prec_w$. For $k \in \{1\}$ we can observe that the choice of $\leqslant_{Lex1}$ has the best sensitivity and the aggregation does not affect the results. The choice of $\leqslant_{Lex2}$ and $\mathcal{A}_{meanpow}$ or $\mathcal{A}_{mean}$ gives very good specificity and precision. For $k \in \{3, 5\}$ we may observe that for $\leqslant_{Lex2}$ and $\mathcal{A}_{meanpow}$ or $\mathcal{A}_{mean}$ we get the best specificity, precision and accuracy, while $\leqslant_{Lex2}$ and $\mathcal{A}_{meanpow}$ or $\mathcal{A}_{mean}$ give us the best sensitivity. For all k the choice of $\mathcal{A}_{meanmax}$ together with $\leqslant_{Lex2}$ gives the lowest accuracy (other measures can also be below the average).
Secondly, we consider $Prec_A$. For all considered k we may observe that the choice of $\leqslant_{Lex1}$ together with $\mathcal{A}_{meanmax}$ gives the best specificity and precision; accuracy is also high. For $k > 1$ the order $\leqslant_2$ always gives the highest sensitivity, however the choice of aggregation is an important aspect. It is worth mentioning here that the aggregation $\mathcal{A}_{mean}$ gives weak results, and therefore it is not included in the following Tables IV-VI.

TABLE I. k=1, $Prec_w$

accuracy	sensitivity	specificity	precision	$S_{\mathcal{A}}$	O
0.941137	0.977688	0.930419	0.960593	$\mathcal{A}_{mean}$	$\leqslant_{XY}$
0.933897	0.964031	**0.940072**	**0.965308**	$\mathcal{A}_{meanpow}$	$\leqslant_{Lex2}$
0.933834	0.964002	**0.940072**	0.965317	$\mathcal{A}_{mean}$	$\leqslant_{Lex2}$
0.920231	0.977667	0.888832	0.937891	$\mathcal{A}_{meanpow}$	$\leqslant_2$
0.916767	0.976582	0.883699	0.935661	$\mathcal{A}_{mean}$	$\leqslant_2$
0.912794	0.990000	0.852067	0.920605	$\mathcal{A}_{meanpow}$	$\leqslant_{XY}$
0.902189	0.985022	0.839038	0.913574	$\mathcal{A}_{meanmax}$	$\leqslant_{XY}$
0.899651	**0.992255**	0.820990	0.905389	$\mathcal{A}_{mean}$	$\leqslant_{Lex1}$
0.899651	**0.992255**	0.820990	0.905389	$\mathcal{A}_{meanpow}$	$\leqslant_{Lex1}$
0.899651	**0.992255**	0.820990	0.905389	$\mathcal{A}_{meanmax}$	$\leqslant_{Lex1}$
0.845358	0.968112	0.741071	0.865637	$\mathcal{A}_{meanmax}$	$\leqslant_2$
0.761087	0.910109	0.614727	0.802463	$\mathcal{A}_{meanmax}$	$\leqslant_{Lex2}$

TABLE II. k=3, $Prec_w$

accuracy	sensitivity	specificity	precision	$S_{\mathcal{A}}$	O
0.927188	0.977216	**0.903770**	**0.946142**	$\mathcal{A}_{mean}$	$\leqslant_{Lex2}$
0.914093	0.966544	0.895268	0.941158	$\mathcal{A}_{meanpow}$	$\leqslant_{Lex2}$
0.908623	0.980723	0.859971	0.923881	$\mathcal{A}_{mean}$	$\leqslant_{XY}$
0.884529	0.976170	0.816262	0.901525	$\mathcal{A}_{meanpow}$	$\leqslant_2$
0.882382	0.982229	0.800397	0.894695	$\mathcal{A}_{meanmax}$	$\leqslant_{XY}$
0.877970	0.981130	0.793262	0.891060	$\mathcal{A}_{mean}$	$\leqslant_2$
0.873287	0.986096	0.773879	0.882572	$\mathcal{A}_{meanmax}$	$\leqslant_{Lex1}$
0.871720	0.972877	0.794825	0.890454	$\mathcal{A}_{meanmax}$	$\leqslant_2$
0.866126	0.988280	0.752465	0.873822	$\mathcal{A}_{meanpow}$	$\leqslant_{XY}$
0.856706	0.988296	0.731050	0.864038	$\mathcal{A}_{meanpow}$	$\leqslant_{Lex1}$
0.850010	**0.990591**	0.711260	0.855724	$\mathcal{A}_{mean}$	$\leqslant_{Lex1}$
0.789597	0.924097	0.672000	0.829244	$\mathcal{A}_{meanmax}$	$\leqslant_{Lex2}$

TABLE III. k=5, $Prec_w$

accuracy	sensitivity	specificity	precision	$S_{\mathcal{A}}$	O
0.896639	0.972264	0.850171	0.917415	$\mathcal{A}_{mean}$	$\leqslant_{Lex2}$
0.893031	0.965679	**0.853832**	**0.918953**	$\mathcal{A}_{meanpow}$	$\leqslant_{Lex2}$
0.886348	0.981766	0.810557	0.899446	$\mathcal{A}_{mean}$	$\leqslant_{XY}$
0.870436	0.970200	0.794800	0.891004	$\mathcal{A}_{meanmax}$	$\leqslant_{XY}$
0.858605	0.972225	0.764412	0.876609	$\mathcal{A}_{meanmax}$	$\leqslant_{Lex1}$
0.856855	0.980619	0.746310	0.869090	$\mathcal{A}_{mean}$	$\leqslant_2$
0.855151	0.951835	0.789845	0.886559	$\mathcal{A}_{meanmax}$	$\leqslant_2$
0.854250	0.968396	0.761289	0.874705	$\mathcal{A}_{meanpow}$	$\leqslant_2$
0.833456	0.979589	0.691257	0.845435	$\mathcal{A}_{meanpow}$	$\leqslant_{XY}$
0.830351	**0.988461**	0.668195	0.836751	$\mathcal{A}_{mean}$	$\leqslant_{Lex1}$
0.825374	0.978993	0.672034	0.837250	$\mathcal{A}_{meanpow}$	$\leqslant_{Lex1}$
0.793329	0.896759	0.728625	0.850618	$\mathcal{A}_{meanmax}$	$\leqslant_{Lex2}$

TABLE IV. k=1, $Prec_A$

accuracy	sensitivity	specificity	precision	$S_{\mathcal{A}}$	O
0.913600	0.972956	0.883822	0.935322	$\mathcal{A}_{meanpow}$	$\leqslant_2$
0.913300	0.959729	**0.905535**	**0.946325**	$\mathcal{A}_{meanpow}$	$\leqslant_{Lex1}$
0.913300	0.959729	**0.905535**	**0.946325**	$\mathcal{A}_{meanmax}$	$\leqslant_{Lex1}$
0.911312	0.962459	0.897076	0.941846	$\mathcal{A}_{meanpow}$	$\leqslant_{XY}$
0.903654	**0.975148**	0.858914	0.923156	$\mathcal{A}_{meanpow}$	$\leqslant_{Lex2}$
0.865831	0.949842	0.818633	0.901164	$\mathcal{A}_{meanmax}$	$\leqslant_2$
0.793718	0.919234	0.691226	0.837084	$\mathcal{A}_{meanmax}$	$\leqslant_{XY}$
0.769863	0.897532	0.661862	0.820662	$\mathcal{A}_{meanmax}$	$\leqslant_{Lex2}$

TABLE V. k=3, $Prec_A$

accuracy	sensitivity	specificity	precision	$S_{\mathcal{A}}$	O
0.924491	0.957286	**0.931892**	**0.960729**	$\mathcal{A}_{meanmax}$	$\leqslant_{Lex1}$
0.902183	0.942547	0.912231	0.949249	$\mathcal{A}_{meanpow}$	$\leqslant_2$
0.902179	0.952512	0.895032	0.940448	$\mathcal{A}_{meanpow}$	$\leqslant_{Lex2}$
0.901164	0.938012	0.917968	0.952032	$\mathcal{A}_{meanpow}$	$\leqslant_{XY}$
0.896202	0.933100	0.916215	0.950736	$\mathcal{A}_{meanpow}$	$\leqslant_{Lex1}$
0.887169	**0.958927**	0.851131	0.917875	$\mathcal{A}_{meanmax}$	$\leqslant_2$
0.840037	0.927569	0.796373	0.887644	$\mathcal{A}_{meanmax}$	$\leqslant_{XY}$
0.824183	0.915723	0.777194	0.877020	$\mathcal{A}_{meanmax}$	$\leqslant_{Lex2}$

TABLE VI. k=5, $Prec_A$

accuracy	sensitivity	specificity	precision	$S_{\mathcal{A}}$	O
0.902454	0.938902	**0.919470**	**0.952262**	$\mathcal{A}_{meanmax}$	$\leqslant_{Lex1}$
0.879100	0.933710	0.878334	0.930097	$\mathcal{A}_{meanpow}$	$\leqslant_{Lex2}$
0.869764	0.899224	0.915010	0.948245	$\mathcal{A}_{meanpow}$	$\leqslant_{XY}$
0.868476	**0.941688**	0.839992	0.910364	$\mathcal{A}_{meanmax}$	$\leqslant_2$
0.866045	0.909146	0.888279	0.934200	$\mathcal{A}_{meanpow}$	$\leqslant_2$
0.864823	0.933260	0.845293	0.912846	$\mathcal{A}_{meanmax}$	$\leqslant_{XY}$
0.857555	0.885251	0.909973	0.944650	$\mathcal{A}_{meanpow}$	$\leqslant_{Lex1}$
0.841804	0.919378	0.814287	0.896339	$\mathcal{A}_{meanmax}$	$\leqslant_{Lex2}$

V. CONCLUSIONS

In this paper, we discussed possible axiomatic definitions of inclusion and similarity measures for interval-valued fuzzy setting and introduced such concepts with widths of intervals involved. Some general formula of generating the similarity between interval-valued fuzzy sets have been proposed. The relationships between the similarity measures and the inclusion measures of intervals and aggregation operations have been investigated. Moreover, we applied the similarity measure in the decision making algorithm based on the interval-valued k-NN method, it is a different method than for example in [40], [41]. In the future,

• we will analyze the impact of using different orders in step IV and in some cases of step VII;
• we will analyze the impact of using different aggregations in Step VI and when building a measure of similarity;
• we will examine the effectiveness of our algorithm for data sets with missing values, which for individual attributes a_i will be supplemented for n objects o, as follows:

$$[\min_{k\in\{1,\ldots,n\}} a_{i_{o_k}}, \max_{k\in\{1,\ldots,n\}} a_{i_{o_k}}];$$

• we will compare the effectiveness of our algorithm with other methods (fuzzy kNN, IV-kNN based on the distance and similarity irrespective of uncertainty);
• we will check whether and to what extent the conclusions obtained are affected by a small modification of the input parameters, e.g. the weak ignorance function or its parameters.

Acknowledgements

This work was partially supported by the Centre for Innovation and Transfer of Natural Sciences and Engineering Knowledge of University of Rzeszów, Poland, the project RPPK.01.03.00-18-001/10.

References

[1] L.A. Zadeh, "Fuzzy sets", Inf. Contr., 8, pp. 338 - 353, 1965.
[2] 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., 24(1), pp. 174 - 194, 2016.
[3] U. Bodenhofer, B. De Baets, J. Fodor, "A compendium of fuzzy weak orders: Representations and constructions", Fuzzy Sets Syst. 158, pp. 811 - 829, 2007.
[4] K. Bosteels, E.E. Kerre, "On a reflexivity-preserving family of cardinality-based fuzzy comparison measures", Inf. Sci., 179, pp. 2342 - 2352, 2009.
[5] H. Bustince, E. Barrenechea, M. Pagola, "Image thresholding using restricted equivalence functions and maximizing the measures of similarity", Fuzzy Sets Syst., 158, pp. 496 - 516, 2007.
[6] B. De Baets, H. De Meyer, H. Naessens, "A class of rational cardinality-based similarity measures", J. Comp. Appl. Math., 132, pp. 51 - 69, 2001.
[7] B. De Baets, S. Janssens, H. De Meyer, "On the transitivity of a parametric family of cardinality-based similarity measures", Internat. J. Appr. Reas., 50, pp. 104 - 116, 2009.
[8] B. Jayaram, R. Mesiar, "I-Fuzzy equivalence relations and I-fuzzy partitions", Inf. Sci., 179, pp. 1278 - 1297, 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., 441, pp. 95 - 112, 2018.
[10] L.A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning", Inf. Sci., 8, Part I, pp. 199 - 251, Part II, pp. 301 - 357, Inf. Sci. 9, Part III, pp. 43 - 80, 1975.
[11] R. Sambuc, "Fonctions ϕ-floues: Application á l'aide au diagnostic en pathologie thyroidienne", Ph.D. Thesis, Universit́e de Marseille, France (in French), 1975.
[12] I.B. Turksen, "Interval-valued fuzzy sets based on normal forms", Fuzzy Sets Syst., 20, pp. 191–210, 1986.
[13] M.B. Gorzalczany, "A method of inference in approximate reasoning based on interval-valued fuzzy sets", Fuzzy Sets Syst., 21, pp. 1-17, 1987.
[14] H. Bustince, J. Fernandez, A. Kolesárová, R. Mesiar, "Generation of linear orders for intervals by means of aggregation functions", Fuzzy Sets Syst., 220, pp. 69 - 77, 2013.
[15] M.J. Asiain, H. Bustince, B. Bedregal, Z. Takáč, M. Baczyński, D. Paternain, G. Dimuro, "About the use of admissible order for defining implication operators", In: J.P. Carvalho et al. (eds.): Springer, IPMU 2016, Part I, CCIS 610, pp. 353 - 362, 2016.
[16] 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., 88, pp. 91 - 109, 2017.
[17] Z.S. Xu, R.R.Yager, "Some geometric aggregation operators based on intuitionistic fuzzy sets", Int. J. Gen. Syst., 35, pp. 417-433, 2006.
[18] G. Beliakov, H. Bustince, T. Calvo, "A practical Guide to Averaging Functions", In: Studies in Fuzziness and Soft Computing. Springer Vol. 329, 2016.
[19] M. Komorníková, R. Mesiar, "Aggregation functions on bounded partially ordered sets and their classification", Fuzzy Sets Syst., 175, pp. 48 - 56, 2011.
[20] G. Deschrijver, E.E. Kerre, "Implicators based on binary aggregation operators in interval-valued fuzzy set theory", Fuzzy Sets Syst., 153(2), pp. 229 - 248, 2005.
[21] 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.
[22] M. Elkano, J.A. Sanz, M. Galar, B. Pękala, U. Bentkowska, H. Bustince, "Composition of interval-valued fuzzy relations using aggregation functions", Inf. Sci., 369, pp. 690 - 703, 2016.
[23] B. Pękala, "Uncertainty Data in Interval-Valued Fuzzy Set Theory. Properties, algirithms and applications", Studies in Fuzziness and Soft Computing, Springer, 2019.
[24] 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., 23(3), pp. 137 - 209, 2000.
[25] W.Y. Zeng, P. Guo, "Normalized distance, similarity measure, inclusion measure and entropy of interval-valued fuzzy sets and their relationship", Inf. Sci., 178, pp. 1334 - 1342, 2008.
[26] 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., 432, pp. 97 - 115, 2018.
[27] 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., 3, pp. 167–177, 2020.
[28] 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
[29] G. Deschrijver, C. Cornelis, "Represantability in interval-valued fuzzy set theory", International Journal of Uncertainty. Fuzziness Knowledge-Based Systems, 15(3), pp. 345–361, 2007.
[30] B. Pękala, U. Bentkowska, J. Fernandez and H. Bustince, "Equivalence measures for Atanassov intuitionistic fuzzy setting used to algorithm of image processing", IEEE Conference on Fuzzy Systems (FUZZIEEE'2019), June 23-26, 2019, New Orleans, Louisiana, USA, 2019.
[31] 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 Sets and Systems, in press
[32] T. M. Cover and P. E. Hart, "Nearest neighbor pattern classification", IEEE Transactions on Information Theory 13(1), 21–27, 1967.
[33] X. Wu, V. Kumar, "The Top Ten Algorithms in Data Mining", Data Mining and Knowledge Discovery. Chapman and Hall/CRC, 2009.
[34] G. Shakhnarovich, T. Darrell, P. Indyk, "Nearest-Neighbor Methods in Learning and Vision: Theory and Practice", The MIT Press, 2006.
[35] G. Bhattacharya, K. Ghosh, A. S. Chowdhury, "An affinity-based new local distance function and similarity measure for kNN algorithm", Pattern Recognition Letters, 33, pp. 356–363, 2012.
[36] A. Stachowiak, P. Żywica, K. Dyczkowski, A. Wójtowicz, "An Interval-Valued Fuzzy Classifier Based on an Uncertainty-Aware Similarity Measure", P. Angelov et al. (eds.), Intelligent Systems'2014, Advances in Intelligent Systems and Computing, 322, Springer, pp. 741 – 751, 2015.
[37] J. Derrac, F. Chiclana, S. García, F. Herrera, "An Interval Valued K-Nearest Neighbors Classifier", 16th World Congress of the International Fuzzy Systems Association (IFSA) and 9th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT), Atlantis Press, pp. 378–384, 2015.
[38] 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.
[39] https://archive.ics.uci.edu/ml/index.php
[40] U. Bentkowska, J.G. Bazan, W. Rząsa, L. Zaręba, "Application of interval-valued aggregation to optimization problem of k-NN classifiers for missing values case", Information Sciences, 486, pp. 434–449, 2019.
[41] U. Bentkowska, "Interval-valued methods in Classifications and Decisions", Studies in Fuzziness and Soft Computing, vol. 378, Springer, 2020.