Aggregation of Fuzzy Equivalences in Data Exploration by kNN Classifier

1st Anna Król
College of Natural Sciences
University of Rzeszów
Rzeszów, Poland
annakrol@ur.edu.pl

2nd Wojciech Rząsa
College of Natural Sciences
University of Rzeszów
Rzeszów, Poland
wrzasa@ur.edu.pl

3rd Piotr Grochowalski
College of Natural Sciences
University of Rzeszów
Rzeszów, Poland
piotrg@ur.edu.pl

***Abstract*—The article deals with notions of fuzzy equivalences connectives, dependencies between them, and their applicability to real-life problem as engineering efficient classifiers. It is a fundamental problem in supervised learning - part of artificial intelligence. In the contribution the problem of preservation of their properties in the aggregation process is considered. Moreover, suitability of some fuzzy equivalences and aggregations are proved in the task of data classification by a k-nearest neighbour algorithm.**

***Index Terms*—aggregation function, fuzzy connective, fuzzy equivalence, fuzzy C–equivalence, k-nearest neighbour algorithm**

I. Introduction

Fuzzy equivalences are known both as a fuzzy relation (e.g. [1], [2]) or a fuzzy connective (e.g. [3]). In the paper diverse approaches to the definition of fuzzy equivalences as a fuzzy connective are considered. We recall two of them that origin from fuzzy algebra and propose a new approach which is related to the concept of measurement of closeness. The possibility of using fuzzy equivalences as closeness measure justifies their application to some problems of data mining, where distance-based algorithms are applied. In this paper we examine a potential of fuzzy equivalences when used in a modified version of k-nearest neighbour classifier. For this purpose, fuzzy equivalences are aggregated by aggregation functions. Many applications of aggregation function to real life problems are known (cf. [4], [5]). The problem of preservation of fuzzy equivalences properties in the aggregation process is considered as well. Moreover, since efficiency of data mining algorithms is a crucial point (both for training and testing stages) we formulate some remarks that allow to reduce number of experiments without loss of information about missed outputs.

In Section II, basic notions, concerning aggregation functions and some fuzzy connectives, useful in the paper are presented. In Section III, practical aspects of fuzzy equivalences and possibilities of their aggregation are described, and in Section IV, the experimental results by the use of practical aggregation of fuzzy equivalences are discussed.

This work was partially supported by the Centre for Innovation and Transfer of Natural Sciences and Engineering Knowledge in Rzeszów, through Project Number RPPK.01.03.00-18-001/10.

II. Preliminaries

A. Aggregation Functions

Now we present useful information about aggregation functions.

Definition 1 (cf. [6], pp. 6-22, [7], pp. 216-218)**.** Let $n \in \mathbb{N}$. A function $A : [0,1]^n \to [0,1]$ which is increasing, i.e. for $x_i, y_i \in [0,1], x_i \leqslant y_i, i = 1, \ldots, n$

$$A(x_1, \ldots, x_n) \leqslant A(y_1, \ldots, y_n)$$

is called an aggregation function if

$$A(0, \ldots, 0) = 0, \ A(1, \ldots, 1) = 1. \quad (1)$$

Moreover, we call an aggregation function A a mean if it is idempotent, i.e.

$$A(x, \ldots, x) = x, \quad x \in [0,1]. \quad (2)$$

Example 1 (cf. [6], pp. 44-56, [8], p. 29)**.** Some examples of aggregation functions are given by standard means such as lattice operations $\min, \max$, projections, geometric mean, harmonic mean and

- weighted arithmetic means

$$A_w(x_1, \ldots, x_n) = \sum_{k=1}^{n} w_k x_k, \quad (3)$$

for $w_k > 0, \sum_{k=1}^{n} w_k = 1$,

- quasi-arithmetic means

$$M_\varphi(x_1, \ldots, x_n) = \varphi^{-1}\Big(\frac{1}{n}\sum_{k=1}^{n} \varphi(x_k)\Big), \quad (4)$$

where $\sum_{k=1}^{n} w_k = 1, x_1, \ldots, x_n \in [0,1]$ and $\varphi : [0,1] \to \mathbb{R}$ is a continuous, strictly increasing function.

B. Fuzzy Conjunctions

Now, the definition and some properties of a fuzzy conjunction is presented.

Definition 2 ([9])**.** An operation $C : [0,1]^2 \to [0,1]$ is called a fuzzy conjunction if it is increasing and

$$C(1,1) = 1, \quad C(0,0) = C(0,1) = C(1,0) = 0.$$

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

Let us observe that fuzzy conjunctions are aggregation functions for $n = 2$. If a binary aggregation function has the zero element $z = 0$ (e.g. geometric mean), then it is a fuzzy conjunction.

Example 2. Consider the following family of fuzzy conjunctions for $a \in [0, 1]$

$$C_a(x, y) = \begin{cases} 1, & \text{if } x = y = 1 \\ 0, & \text{if } x = 0 \text{ or } y = 0 \\ a & \text{otherwise} \end{cases}. \tag{5}$$

Operations C_0 and C_1 are the least and the greatest fuzzy conjunction, respectively.

Example 3. Other examples of fuzzy conjunctions are listed below. Among them we recall the well-known t-norms: minimum, product, Łukasiewicz, drastic, which are denoted in the traditional way T_M, T_P, T_{LK}, T_D, respectively:

$$C_2(x, y) = \begin{cases} y, & \text{if } x = 1 \\ 0, & \text{if } x < 1 \end{cases},$$
$$C_3(x, y) = \begin{cases} x, & \text{if } y = 1 \\ 0, & \text{if } y < 1 \end{cases},$$
$$C_4(x, y) = \begin{cases} 0, & \text{if } x + y \leqslant 1 \\ y, & \text{if } x + y > 1 \end{cases},$$
$$T_M(x, y) = \min(x, y),$$
$$T_P(x, y) = xy,$$
$$T_{LK}(x, y) = \max(x + y - 1, 0),$$
$$T_D(x, y) = \begin{cases} x, & \text{if } y = 1 \\ y, & \text{if } x = 1 \\ 0, & \text{otherwise} \end{cases}.$$

C. Fuzzy Implications

Definition 3 ([10])**.** A binary operation $I\colon [0, 1]^2 \to [0, 1]$ is called a fuzzy implication if it is decreasing with respect to the first variable and increasing with respect to the second variable and

$$I(0, 0) = I(0, 1) = I(1, 1) = 1, \quad I(1, 0) = 0.$$

Definition 4. We say that a fuzzy implication I fulfils the identity principle (IP) if

$$I(x, x) = 1, \quad x \in [0, 1]. \tag{IP}$$

Example 4 ([10], pp. 4,5)**.** The operations I_0 and I_1 are the least and the greatest fuzzy implication, respectively, where

$$I_0(x, y) = \begin{cases} 1, & \text{if } x = 0 \text{ or } \ y = 1 \\ 0, & \text{otherwise} \end{cases},$$
$$I_1(x, y) = \begin{cases} 0, & \text{if } x = 1, y = 0 \\ 1, & \text{otherwise} \end{cases}.$$

The following are the other examples of fuzzy implications.

$$I_{\text{LK}}(x, y) = \min(1 - x + y, 1),$$
$$I_{\text{GD}}(x, y) = \begin{cases} 1, & \text{if } x \le y \\ y, & \text{if } x > y \end{cases},$$
$$I_{\text{RC}}(x, y) = 1 - x + xy,$$
$$I_{\text{DN}}(x, y) = \max(1 - x, y),$$
$$I_{\text{GG}}(x, y) = \begin{cases} 1, & \text{if } x \le y \\ \frac{y}{x}, & \text{if } x > y \end{cases},$$
$$I_{\text{RS}}(x, y) = \begin{cases} 1, & \text{if } x \le y \\ 0, & \text{if } x > y \end{cases},$$
$$I_{\text{FD}}(x, y) = \begin{cases} 1, & \text{if } x \le y \\ \max(1 - x, y), & \text{if } x > y \end{cases},$$
$$I_{\text{WB}}(x, y) = \begin{cases} 1, & \text{if } x \le 1 \\ y, & \text{if } x = 1 \end{cases},$$
$$I_{\text{DP}}(x, y) = \begin{cases} y, & \text{if } x = 1 \\ 1 - x, & \text{if } y = 0 \\ 1, & \text{otherwise} \end{cases}.$$

The implications fulfilling the property IP are: I_1, I_{LK}, I_{GD}, I_{GG}, I_{RS}, I_{WB}, I_{FD}, I_{DP}.

D. Fuzzy equivalences

Let us focus on fuzzy equivalences which can be defined in different ways.

Definition 5 ([11], cf. [12], p. 33)**.** Let C be a fuzzy conjunction. A fuzzy C–equivalence is a function $E : [0, 1]^2 \to [0, 1]$ fulfilling the following conditions for $x, y, z \in [0, 1]$.

$$E(0, 1) = 0 \qquad \text{(boundary property)}, \tag{6}$$

$$E(x, x) = 1, \qquad \text{(reflexivity)}, \tag{7}$$

$$E(x, y) = E(y, x), \qquad \text{(symmetry)}, \tag{8}$$

$$C(E(x, y), E(y, z)) \leqslant E(x, z), \qquad (C\text{-transitivity}). \tag{9}$$

Further generalizations of fuzzy C–equivalence can be found in [11].

A fuzzy equivalence considered by Fodor and Roubens in [3] is defined as follows.

Definition 6 ([3], p. 33)**.** A Fodor-Roubens fuzzy equivalence is a function $E : [0, 1]^2 \to [0, 1]$ which fulfils (6), (7), (8), and

$$E(x, y) \le E(u, v), \ x \le u \le v \le y, \ x, y, u, v \in [0, 1]. \tag{10}$$

There exists a characterization of such defined fuzzy equivalence by the use of fuzzy implications fulfilling (IP).

Theorem 1 ([3], p. 33)**.** *A function $E : [0, 1]^2 \to [0, 1]$ is a fuzzy equivalence if and only if there exists such a fuzzy implication I fulfilling (IP) that*

$$E_I(x, y) = \min(I(x, y), I(y, x)), \ x, y \in [0, 1]. \tag{11}$$

If, in the place of $\min$ in (11) we consider any fuzzy conjunction, then we obtain more generalized definition of a fuzzy equivalence.

Definition 7 ([13]). Let C, I be a fuzzy conjunction and implication, respectively. The function $E : [0,1]^2 \to [0,1]$ given by the formula

$$E_{C,I}(x,y) = C(I(x,y), I(y,x)),\ x,y \in [0,1] \qquad (12)$$

will be called (C, I)-equivalence.

Remark 1. ([13]) The operation $E_{C,I}$ given by (12) fulfils zero-one table of crisp equivalence. Additionally, if I fulfils (IP), then $E_{C,I}$ is reflexive, i.e. it fulfils (7). Moreover, if C is a commutative fuzzy conjunction, then $E_{C,I}$ is symmetric, i.e. it fulfils (8).

Theorem 2 (cf. [3], p.27). *Let T be a left continuous t-norm, and I_T its residual implication. Then the following inequality holds for all $x, y, z \in [0, 1]$*

$$T((I_T(x,y), I_T(y,z)) \leq I_T(x,z).$$

Theorem 3. *Let T be a left continuous t-norm, and I_T its residual implication. Then operation $E_{\min,I_T}$ is a fuzzy C-equivalence for any commutative fuzzy conjunction $C \leq T$.*

Proof. Let us observe that a residual implication I_T fulfils (IP), so according to Remark 1 the operation $E_{\min,I_T}$ fulfils (6), (7), (8). It is enough to show that the operation $E_{\min,I_T}$ is C-transitive. Let $x, y, z \in [0, 1]$. By (12), monotonicity of C, the condition $C \leq T$ and Theorem 2 we obtain as the following

$$C(E_{\min,I_T}(x,y), E_{\min,I_T}(y,z))$$
$$= C(\min(I_T(x,y), I_T(y,x)), \min(I_T(y,z), I_T(z,y)))$$
$$\leq C((I_T(x,y), I_T(y,z)) \leq T((I_T(x,y), I_T(y,z)) \leq I_T(x,z).$$

Similarly we can show that

$$C(E_{\min,I_T}(x,y), E_{\min,I_T}(y,z)) \leq I_T(z,x).$$

Hence

$$C(E_{\min,I_T}(x,y), E_{\min,I_T}(y,z))$$
$$\leq \min(I_T(x,z), I_T(z,x)) = E_{\min,I_T}(x,z).$$

□

III. Practical aspects of fuzzy equivalences and their aggregation

A. Various approaches to the notion of fuzzy equivalence

Let us consider the exclusive disjunction of the classical propositional calculus and its generalization in the fuzzy logic, i.e. the function $\bar{D} : [0,1]^2 \to [0,1]$, fulfilling boundary conditions $\bar{D}(0,0) = \bar{D}(1,1) = 0$, $\bar{D}(0,1) = \bar{D}(1,0) = 1$. This operation can be associated with a discrete metric on a unit interval. In classical logic, the equivalence is the negation of exclusive disjunction. Analogously, in fuzzy logic it can be expressed by duality of these operation, i.e. their values for equal arguments sum up to 1. Continuing this analogy, but in relation to the concept of a metric, we notice that while the metric is a measure of the distance of two points, fuzzy equivalence can be interpreted in a dual way, i.e. as a measure of the closeness of two points. Indeed, crisp equivalence connective has value 1 for arguments with the same logical value, and 0 for arguments with different logical values, so - using the closeness interpretation - it takes value 1 if the closeness of the arguments is 'complete', and value 0 when the closeness of arguments is the smallest possible. As a consequence of the interpretation of fuzzy equivalence as closeness, as in the case of metrics (and pseudo-metrics), it is natural to require closeness to have the property of symmetry.

If we additionally consider the triangle condition $d(x,y) + d(y,z) \leq d(x,z)$ from the definition of a metric (and pseudo-metric), then its equivalent for fuzzy equivalence is $E(x,y) + E(y,z) - 1 \leq E(x,z)$ which can be expressed by Łukasiewicz t-norm T_{LK} as follows: $T_{LK}(E(x,y), E(y,z)) \leq E(x,z)$. Replacing Łukasiewicz t-norm with any fuzzy conjunction C and adding for fuzzy equivalence the requirement that $C(E(x,y), E(y,z)) \leq E(x,z)$ leads to the definition of fuzzy C-equivalence (Definition 5).

Conducting purely cognitive considerations can lead to definitions of fuzzy equivalences by the use of a variety of formulas. However, with regard to some practical applications of fuzzy equivalence connective, as a factor determining the degree of closeness of two points, it seems natural to require that for three values x, y, z in the natural order $x \leq y \leq z$, assume that x and z are not closer to each other than with the element y 'separating' them. This can be expressed formally by the conditions for all $x, y, z \in [0, 1]$ such that $x \leq y \leq z$

$$E(x,z) \leq E(y,z) \quad \text{and} \quad E(x,z) \leq E(x,y). \qquad (13)$$

Condition (13) is equivalent to (10) (what was shown in [14]), which leads to the definition of fuzzy equivalence introduced by Fodor and Roubens (Definition 6).

Now, let us focus on interpretations of selected fuzzy equivalences.

Example 5. Table I presents examples of fuzzy equivalences generated by the use of formula (11) and these of the fuzzy implications from Example 4 that fulfil (IP).

Assessing the chosen equivalences from the perspective of their use in practical applications as a closeness measure, it can be seen that fuzzy equivalences E_{RS} and E_1 may be applied to problems and phenomena described by the values from a symbolic scale of measure, E_{GD} and E_{WB} can be applied for ordinal scale, E_{LK}, E_{FD} and E_{DP} when the values are from an interval scale, and finally E_{GG} in the case where the values are from a ratio scale.

All the above formulas take into account only two points' values (namely x and y) in the assessment of their closeness. We will call this type of closeness absolute. However, it is possible to define the closeness of two points, taking into account other points in the problem to which the closeness factor is applied. In this sense, such closeness will be called relative.

TABLE I
FUZZY EQUIVALENCES

I	E_I
I_{LK}	$E_{LK}(x,y) = 1 - \|x-y\|$
I_{GD}	$E_{GD}(x,y) = \begin{cases} 1, & \text{if } x = y \\ x, & \text{if } x < y \\ y, & \text{if } x > y \end{cases}$
I_{GG}	$E_{GG}(x,y) = \begin{cases} 1, & \text{if } x = y \\ \frac{x}{y}, & \text{if } x < y \\ \frac{y}{x}, & \text{if } x > y \end{cases}$
I_{RS}	$E_{RS}(x,y) = \begin{cases} 1, & \text{if } x = y \\ 0, & \text{if } x \neq y \end{cases}$
I_{WB}	$E_{WB}(x,y) = \begin{cases} 1, & \text{if } x \neq 1,\ y \neq 1 \\ x, & \text{if } y = 1 \\ y, & \text{if } x = 1 \end{cases}$
I_{FD}	$E_{FD}(x,y) = \begin{cases} 1, & \text{if } x = y \\ \max(1-y, x), & \text{if } x < y \\ \max(1-x, y), & \text{if } x > y \end{cases}$
I_{DP}	$E_{DP}(x,y) = \begin{cases} x, & \text{if } y = 1 \\ y, & \text{if } x = 1 \\ 1-x, & \text{if } y = 0 \\ 1-y, & \text{if } x = 0 \\ 1 & \text{otherwise} \end{cases}$
I_1	$E_1(x,y) = \begin{cases} 0, & \text{if } \{x,y\} = \{0,1\} \\ 1 & \text{otherwise} \end{cases}$

Let U be a universe of a discrete problem with card(U) ≤ 3 called a set of cases, which are described by one feature a with values from $[0,1]$. Since $a : U \to [0,1]$ then the value of feature a for the case $u \in U$ will be denoted by $a(u)$. The closeness of any pair of cases $x, y \in U$ can be expressed by means of closeness of $a(x)$ and $a(y)$ as follows:

$$\begin{cases} 1, & x = y \\ \frac{card(\{\{u,v\}: u,v \in U \wedge \{u,v\} \neq \{x,y\} \wedge E(a(u),a(v)) \leq E(a(x),a(y))\})}{card(\{\{u,v\}: u,v \in U \wedge \{u,v\} \neq \{x,y\}\})}, & otherwise \end{cases} \tag{14}$$

where E is any closeness of absolute type. It contains information about what part of all pairs of cases from the set U is at most as close to each other as the pair (x, y).

It is easy to see that the above method of measuring the closeness of cases may not meet the boundary condition. This happens when the cases x, y such that $a(x) = 0$, $a(y) = 1$ are not the only one fulfilling property $E(a(u), a(v)) = 0$. On the other hand, the above formula meets the conditions of reflexivity and symmetry.

Now, let us consider the set $U = [0,1]$. Let us denote $R(x,y) = \{(u,v) \in U^2 : E(u,v) \leq E(x,y)\}$ and $E^R(x,y)$ denotes the field of the figure $R(x,y)$. If the absolute closeness of $x, y \in [0,1]$ is expressed by the formula E_{LK}, then the corresponding relative closeness of $E^R_{LK}(x,y)$ is the sum of the fields of two congruent isosceles right triangles with side length of $E_{LK}(x,y))$, denoted on Fig. 1 as *e* (see Fig. 1). So, in this case

$$E^R_{LK}(x,y) = (E_{LK}(x,y))^2. \tag{15}$$

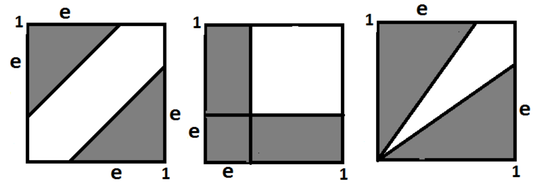

Fig. 1. A picture of $E^R_{LK}(x,y)$,$E^R_{GD}(x,y)$, and $E^R_{GG}(x,y)$

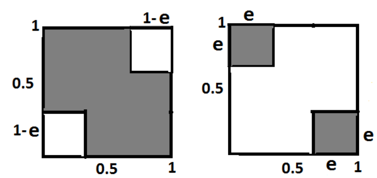

Fig. 2. A picture of $E^R_{FD}(x,y)$

If the absolute closeness of $x, y \in [0,1]$ is expressed by the formula E_{GD}, then the corresponding relative closeness of $E^R_{GD}(x,y)$ is the difference between the square field with side 1 and the square field with side $1 - E_{GD}(x,y)$. The value $E_{GD}(x,y)$ is denoted on figure 1 as *e*. Therefore, in this case

$$E^R_{GD}(x,y) = E_{GD}(x,y) * (2 - E_{GD}(x,y)). \tag{16}$$

If the absolute closeness of $x, y \in [0,1]$ is expressed by the formula E_{GG}, then the corresponding relative closeness of $E^R_{GG}(x,y)$ is the sum of the fields of two congruent right triangles with side lengths of 1 and $E_{GG}(x,y)$, denoted on figure 1 as *e*. Therefore, in this case

$$E^R_{GG}(x,y) = E_{GG}(x,y). \tag{17}$$

If the absolute closeness of $x, y \in [0,1]$ is expressed by the formula E_{FD}, then the corresponding relative closeness of $E^R_{FD}(x,y)$ is the sum of the fields of two squares with side length $E_{FD}(x,y)$ or the difference between the square field with side 1 and the sum of the fields of two squares with side length $1 - E_{FD}(x,y)$, where $E_{FD}(x,y)$ is denoted on figure 2 as *e* (see Fig. 2). Therefore, in this case

$$E^R_{FD}(x,y) = \begin{cases} 1 - 2 \cdot (1 - E_{FD}(x,y))^2, \text{if } E_{FD}(x,y) \geq 0.5 \\ 2 \cdot (E_{FD}(x,y))^2, \quad \text{otherwise} \end{cases} \tag{18}$$

Remark 2. Let $U = [0,1]$ and E be Fodor-Roubens fuzzy equivalence then the corresponding $E^R(x,y)$ is Fodor-Roubens fuzzy equivalence.

Now, let us recall a notion of ordinal equivalence introduced in [15].

Definition 8 (cf. [15])**.** Let $E, F : [0,1]^2 \to [0,1]$ be fuzzy equivalences. We say that E and F are orderly equivalent, and denote $E \backsim F$, if

$$E(x,y) < E(u,v) \Leftrightarrow F(x,y) < F(u,v), x,y,u,v \in [0,1]. \tag{19}$$

Proposition 1 (cf. [15])**.** *The relation given by* (19) *is an equivalence relation.*

Proposition 2 (cf. [15])**.** *Let $E \backsim F$. Then*

$$E(x,y) = E(u,v) \Leftrightarrow F(x,y) = F(u,v), x,y,u,v \in [0,1]. \tag{20}$$

By the properties of monotonic function we obtain the following observation.

Proposition 3. *Let $g : [0,1] \to [0,1]$ be an increasing function, and E, F be fuzzy equivalences. If $E(x,y) = g(F(x,y))$, then $E \backsim F$.*

Regarding to the four considered fuzzy equivalences we can observe the following facts.

Remark 3. Fuzzy equivalences $E_{LK}, E_{GD}, E_{GG}, E_{FD}$ are orderly equivalent to their corresponding equivalences $E^R_{LK}, E^R_{GD}, E^R_{GG}, E^R_{FD}$.
It results from proposition 3 and the fact that each of the fuzzy E^R equivalences is an increasing bijection on $[0,1]$ of the corresponding fuzzy equivalence E.
For instance for the pair E_{LK}, E^R_{LK}, and any $x,y,u,v \in [0,1]$ we have

$$E_{LK}(x,y) < E_{LK}(u,v) \Leftrightarrow (E_{LK}(x,y))^2 < (E_{LK}(u,v))^2.$$

Remark 4. None two of the fuzzy equivalences $E_{LK}, E_{GD}, E_{GG}, E_{FD}$ are orderly equivalent.
For example, let us observe that for $(x,y) = (0.1, 0.3)$ and $(u,v) = (0.3, 0.6)$ we have on one hand

$$E_{LK}(0.1, 0.3) = 0.8 > 0.7 = E_{LK}(0.3, 0.6),$$
$$E_{FD}(0.1, 0.3) = 0.7 > 0.4 = E_{FD}(0.3, 0.6).$$

On the other hand

$$E_{GD}(0.1, 0.3) = 0.1 < 0.3 = E_{GD}(0.3, 0.6),$$
$$E_{GG}(0.1, 0.3) = 0.(3) < 0.5 = E_{GG}(0.3, 0.6).$$

It means that none of E_{LK}, E_{FD} is orderly equivalent with any of E_{GD}, E_{GG}.

Moreover for $(x,y) = (0.9, 0.5)$ and $(u,v) = (0.5, 0.4)$ we may observe in analogous way that E_{LK} is not orderly equivalent E_{FD} as well as E_{GD} is not with E_{GG}.

From the above two remarks we can conclude the following.

Remark 5. None two of the fuzzy equivalences $E^R_{LK}, E^R_{GD}, E^R_{GG}, E^R_{FD}$ are orderly equivalent.

The ordinal equivalence relation can be used to compare fuzzy equivalences with (pseudo) metrics specified on $[0,1]$. Going back to the observation from the introductory part of the work that (pseudo) metrics and equivalences are dual in some sense we propose the following definition.

Definition 9. Let $E, D : [0,1]^2 \to [0,1]$, E be a fuzzy equivalence, D be a metric. We say that E and D are reverse orderly equivalent, and denote $E \simeq F$, if

$$E(x,y) < E(u,v) \Leftrightarrow D(x,y) > D(u,v), x,y,u,v \in [0,1]. \tag{21}$$

Proposition 4. *Let $g : [0,1] \to [0,1]$ be a decreasing function, E be a fuzzy equivalence and D be a metric. If $E(x,y) = g(D(x,y))$, then $E \simeq D$.*

Example 6. Let us observe that fuzzy equivalence E_{LK} and absolute value metric $D = |x-y|$, for $x, y \in [0,1]$ are reverse orderly equivalent.

B. Aggregation of fuzzy equivalences

Properties of aggregation for fuzzy relations were presented In [3]. Here we consider aggregation of fuzzy equivalences defined previously.

Definition 10 (cf. [3], p. 14)**.** Let $n \in \mathbb{N}$ and $A : [0,1]^n \to [0,1]$ be an arbitrary aggregation function. For given fuzzy equivalences $E_1, \ldots, E_n$, we consider a binary operation for all $x, y \in [0,1]$

$$E(x,y) = A(E_1(x,y), \ldots, E_n(x,y)). \tag{22}$$

We say that a class of fuzzy equivalences is closed under an aggregation function A if the result of aggregation belongs to this class for arbitrary fuzzy equivalences from the class.

Theorem 4 (cf. [13])**.** *The family of all Fodor-Roubens fuzzy equivalences is closed under any aggregation function A.*

Theorem 5 ([16])**.** *The family of all fuzzy C-equivalences is closed under aggregation functions A that dominate C ($A \gg C$), i.e.*

$$A(C(a_{1,1}, a_{1,2}), \ldots, C(a_{n,1}, a_{n,2})) \geqslant$$
$$\geqslant C(A(a_{1,1}, \ldots, a_{n,1}), A(a_{1,2}, \ldots, a_{n,2})).$$

Example 7. Minimum preserves fuzzy C-equivalence for any fuzzy conjunction C. Weighted minimum preserves fuzzy C-equivalence for any t-seminorm C.

Definition 11. Let $n \in \mathbb{N}$ and A be any aggregation function. For given fuzzy equivalence E, we consider a binary operation for all $\mathbf{x}, \mathbf{y} \in [0,1]^n$

$$A(E)(\mathbf{x}, \mathbf{y}) = A(E(x_1, y_1), \ldots, E(x_n, y_n)). \tag{23}$$

Definition 12. Let $n \in \mathbb{N}$ and $A1, A2 : [0,1]^n \to [0,1]$ be any aggregation functions, $E, F : [0,1]^2 \to [0,1]$ - fuzzy equivalences, $D : [0,1]^2 \to [0,1]$ a metric. We say that $A1(E)$ is orderly equivalent to $A2(F)$ and denote $A1(E) \backsim A2(F)$ if for all $\mathbf{x}, \mathbf{y}, \mathbf{u}, \mathbf{v} \in [0,1]^n$

$$A1(E)(\mathbf{x}, \mathbf{y}) < A1(E)(\mathbf{u}, \mathbf{v}) \Leftrightarrow A2(F)(\mathbf{x}, \mathbf{y}) < A2(F)(\mathbf{u}, \mathbf{v}). \tag{24}$$

We say tat $A1(E)$ is orderly equivalent to $A2(D)$ and denote $A1(E) \simeq A2(D)$ if for all $\mathbf{x}, \mathbf{y}, \mathbf{u}, \mathbf{v} \in [0,1]^n$

$$A1(E)(\mathbf{x}, \mathbf{y}) < A1(E)(\mathbf{u}, \mathbf{v}) \Leftrightarrow A2(D)(\mathbf{x}, \mathbf{y}) > A2(D)(\mathbf{u}, \mathbf{v}). \tag{25}$$

Proposition 5 (cf. [15])**.** *Ordinal equivalences given by* (24) *and* (25) *are equivalence relations.*

By the properties of monotonic function we obtain the following observations.

Proposition 6. *Let $g : [0,1] \to [0,1]$ be an increasing function, $A1, A2$ be aggregation functions, E, F - fuzzy equivalences. If $A1(E)(\mathbf{x},\mathbf{y}) = g(A2(F)(\mathbf{x},\mathbf{y}))$, then $A1(E) \backsim A2(F)$.*

Proposition 7. *Let $g : [0,1] \to [0,1]$ be a decreasing function, $A1, A2$ be aggregation functions, E a fuzzy equivalence and D. If $A1(E)(\mathbf{x},\mathbf{y}) = g(A2(D)(\mathbf{x},\mathbf{y}))$, then $A1(E) \simeq A2(D)$.*

Now, let us consider the following aggregation functions:
$A1$-arithmetic mean,
$A2_p$-power-root mean, where $p > 0$,
$A3$ - minimum,
$A4$ - maximum,
and a metric $D(x,y) = |x-y|$ for $x, y \in [0,1]$.
Let us recall that we have $A2_p(D)$ - Minkowski metric for $p > 0$, $A2_2(D)$ - Euclidean metric, $A2_1(D)$ - Manhattan metric, $A4(D)$ - Chebyshev metric.

Corollary 1. $A1(E_{LK}) \simeq A2_1(D)$.

Proof. Let us observe that

$$A1(E_{LK})(\mathbf{x},\mathbf{y}) = \frac{1}{n} \cdot \sum_{i=1}^{n}(1 - |x_i - y_i|) =$$

$$= 1 - \frac{1}{n} \cdot \sum_{i=1}^{n}(|x_i - y_i|) = 1 - \frac{1}{n} \cdot A2_1(D)(\mathbf{x},\mathbf{y}).$$

Using Proposition 7 with $g(x) = 1 - \frac{1}{n} \cdot x$ ends the proof. □

Theorem 6. *Let E, F be fuzzy equivalences such that $E \backsim F$. Then $A3(E) \backsim A3(F)$ and $A4(E) \backsim A4(F)$.*

Corollary 2. *For any fuzzy equivalence E from the set $\{E_{LK}, E_{GD}, E_{GG}, E_{FD}\}$ and corresponding fuzzy equivalence E^R from the set $\{E^R_{LK}, E^R_{GD}, E^R_{GG}, E^R_{FD}\}$ we have $A3(E) \backsim A3(E^R)$ and $A4(E) \backsim A4(E^R)$.*

IV. kNN CLASSIFIER WITH AGGREGATIONS OF FUZZY EQUIVALENCES

A. kNN algorithm

One of the most popular machine learning algorithms is the k nearest neighbours algorithm (for short kNN (algorithm)). It represents the class of supervised learning algorithms. The purpose of algorithms from this class is to learn from historical data how values of the features describing input cases determine belonging these cases to known categories. Then, supervised learning algorithms can use this knowledge in the process of qualifying new cases with unknown category membership to proper category. The detailed specification of the k nearest neighbours algorithm is as follows:
Input
Dec - set of categories
Tr - training (historical) set of objects described by a set of attributes A, with known membership to one of categories from Dec
Ts - set of objects described by a set of attributes A, with unknown membership to categories from Dec
k - number of nearest neighbors, k - natural number not greater than cardinality of Tr
$Dist$ - distance measure defined in the space generated by ranges of values of attributes from set A
Output
Membership of each object from set Ts to one of categories from Dec.
Algorithm
1. Repeat for each object $o \in Ts$;
2. Among the objects $u \in Tr$ find k ones that have the closest distance to the object o;
3. Return the category most often represented by k objects designated in the previous step.

The above description reflects well the main idea of the kNN method. However, to be implemented, it needs more detailed description. The selection criterion for k objects from the set Ts (shown in the step 2) is insufficient because it may happen that more than k objects will be at the same distance. In this situation, we need to select as few as possible, but not less than k objects with the smallest possible distances from the object o. The description in the step 3 also needs to be clarified as it may happen that more than one category occurs the maximum number of times. In practice, various approaches are used to resolve such ambiguity. One of them is the assumption that each of the closest neighbours "votes" for this category for the object o from the set Ts, which it represents itself with the weight inversely proportional to its distance to the object o. In this approach, the ambiguity of the criterion for choosing the category for object o (step 3) is a sporadic situation. For the purposes of this study, the following solution was applied: if more than one category occurs most frequently, the one that appears earlier in the input data set is proposed (a pseudo-random approach).

The main goal of this work is to verify the suitability of using fuzzy equivalences in the kNN algorithm. More precisely, we aim to comparing the quality of the classification made using the kNN algorithm and selected metrics with the quality of the classification using the same algorithm when diverse aggregations of fuzzy equivalences are used instead of the metrics. To do this, in step 2 of the algorithm, instead of a metric a function $1 - A(E)$ was used, where A and E are an aggregation function and a fuzzy equivalence, respectively. Let us pay attention to another aspect of the kNN algorithm's operation. In the process of selecting k nearest neighbors, the involved metric or aggregation of fuzzy equivalence de facto acts as a rank operator. This means that it is more important to order the objects of the Tr set with respect to their distance from the object $o \in T$s (and choosing k among them with the smallest distance from o) than using exact value of this distance. The usefulness of Propositions 6, 7 and Theorem 6 can be seen here. Any aggregation of fuzzy equivalences or a metric can be replaced with another, orderly equivalent to the given, and the selection of k nearest neighbors will not change. As a consequence, we can reduce the number of operations performed by the computer when choosing k nearest neighbors and restrict the computations for this step of

TABLE II
DESCRIPTION OF DATA SETS

ID and data set name	data size	number of categories
1. Banknote authentication	1372×5	2
2. QSAR biodegradation	1055×41	2
3. Diabetic retinopathy Debrecen	1151×20	2
4. Fertility diagnosis	100×10	2
5. German credit data	1000×20	2
6. Iris	150×5	3
7. Parkinson speech (train)	1040×26	2
8. Spambase	4601×57	2
9. Wine quality - red	1599×12	6
10. Zoology	101×17	10

the kNN algorithm for one of the orderly equivalent operation (compare Proposition 1 and Corollary 2) .

B. Description of experiments

For the experiments, 10 data sets from the UCI ML repository [17] were used. Their synthetic description is presented in Table II. Implementations of the kNN algorithm in Python, available in the scikit-learn library were exploited.

The following metrics, aggregations and fuzzy equivalences were used to determine k nearest neighbours: three metrics $A4(D)$, $A2_2(D)$, $A2_1(D)$, four aggregation functions $A1$, $A2$, $A3$, $A4$ including aggregation $A2$ with parameter $p \in \{0.5, 2, 3, 4\}$ used with respect to eight fuzzy equivalences $E_{LK}, E_{GD}, E_{GG}, E_{FD}$, $E^R_{LK}, E^R_{GD}, E^R_{GG}, E^R_{FD}$. The choice of parameter p values was partly inspired by [4]. During the experiments, it was benefited that $E^R_{GG} = E_{GG}$ and the relationships expressed in corrolaries 1 and 2.

For each data set from Table II the kNN algorithm was used with $k = 3, 5, 10, 20, 30$, three metrics and 42 compositions of aggregations and fuzzy equivalences (instead of 56 if mentioned corollaries not engaged). Each data set was tenfold divided into training and test parts in a $9 : 1$ ratio using 10-fold cross validation (CV) technique. Classification quality was determined using the accuracy coefficient. It defines, what part of the test objects the classifier correctly assigned to individual category. It was the basic measure that was used when analyzing the results of experiments and formulating conclusions. All following observation concern the specified above set of parameters.

At the beginning the influence of parameters k, Aj, E_i to output classification accuracy was examined. It was checked, which of these 3 parameters affects accuracy the most. For this purpose, separately for each data set, average values of accuracy were computed:
(a) for each k and all $A(E)$ and 10 iterations of CV,
(b) for each A and all k, E and 10 iterations of CV,
(c) for each E and all k, A and 10 iterations of CV.
After that intervals of $[\min, \max]$ values from these computed for (a), (b), (c) were determined, as well as parameter p values corresponding to $\min$ and $\max$ values. The results are presented in Table III. General observation is that for all data sets, excluding 3rd and 10th ones least difference between $\max$ and $\min$ is for parameter k. On the other hand, for all data sets excluding the 3rd one biggest difference is for aggregation. Wilcoxon matched pairs tests implemented in Statistica software [18] used for each pair (k, A), (k, E), (A, E) showed that with significance parameter $p = 0.125$ appropriate choice of aggregations influence output accuracy the most and the influence of parameter k on output is the weakest. Second observation is that aggregation A1 and A2 gives accuracy very better than A4 and most of all than A3 (see column 3 in Table III).

Next, we examined, if little difference in closeness measurement on each attribute (received by using different fuzzy equivalences) may lead to significantly different classification accuracy. Now, we focused on comparing outputs for E_i equivalences and their counterparts from the family of E^R_i equivalences. It is because an easy observation of E^R_i formulas shows that for E^R_{LK} and E^R_{GD} differs from E_{LK} and E_{GD}, respectively on 0.25 at most (only when $E_{LK}(x, y) = 0.5$ and $E_{GD}(x, y) = 0.5$) and E^R_{FD} differs from E_{FD} on at most 0.125 (only when $E_{FD}(x, y) = 0.25$ or $E_{FD}(x, y) = 0.75$). Experiments showed that for more than 99% cases the difference in output accuracy was between -0.1 and 0.1. However, for 0.9% cases the difference was at least 0.2 and twice happened (for $A4(E_{LK})$ of course for $A4(E^R_{LK})$) that the difference was over 0.4. For 18% cases, family of E^R equivalences gave better accuracy than corresponding E and for $16, 5\%$ cases the converse was true. For over 65% the output accuracy was the same.

Next, output accuracy received for each aggregation and fuzzy equivalence was compared with output accuracy received for reference, widely used distance metrics, such that Chebyshev, Euclidean, Manhattan ones. Table IV presents outputs for those 3 distance metrics and aggregations of fuzzy equivalences with 3 best accuracy values. Notice, that on the basis of formula E^R_{GG} and Corollary 1 outputs for $A(E^R_{GG})$ and $A1(E_{LK})$ are the same as for $A(E_{GG})$ and Manhattan metric, respectively. Therefore, they are not presented explicitly, even if they are one of 3 best values for aggregations.

V. CONCLUSIONS

In the paper diverse approaches to the definition of fuzzy equivalences as a fuzzy connective were considered. Some dependencies between them were indicated. However, it seems interesting to examine further relationships as well as usefulness of the approaches in applications.
Practical utility of mean and power-root aggregations as well as some fuzzy Fodor-Roubens equivalences (especially E_{LK}, E^R_{LK}, E_{GG}) in data classification tasks were observed during experiments of real-life data sets. Compositions of theirs or with the use of other means (like geometric, harmonic, OWA ones) may be an interesting alternative for classical metrics in application to distance-based classifiers. On the other hand, $\min$ and $\max$ aggregations seem to be incidentally useful in this case. There are still several aspects of using aggregations of fuzzy equivalences to be examined during experiments which may result with valuable observations useful in data

TABLE III
CLASSIFICATION ACCURACY

data set	[min acc(k), max acc(k)]	[min acc(Aj), max acc(Aj)]	[min acc(E), max acc(E)]
1	[0.796, 0.807]; [30, 3]	[0.588, 0.882]; [A4, A1=A2]	[0.605, 0.892]; [E_{GD}, E^R_{LK}]
2	[0.679, 0.705]; [30, 3]	[0.535, 0.844]; [A3, A1]	[0.668, 0.708]; [E_{LK}, E^R_{FD}]
3	[0.879, 0.890]; [5, 20=30]	[0.886, 0.890]; [A2, A1]	[0.880, 0.892]; [E^R_{GD}, E_{GG}]
4	[0.463, 0.506]; [30, 3]	[0.326, 0.715]; [A3, A1]	[0.453, 0.504]; [E^R_{GG}, E_{FD}]
5	[0.818, 0.841]; [5, 20]	[0.642, 0.874]; [A4, A2]	[0.793, 0.866]; [E^R_{GD}, E^R_{LK}]
6	[0.517, 0.534]; [30, 5]	[0.463, 0.615]; [A3, A1]	[0.510, 0.542]; [E^R_{FD}, E_{LK}]
7	[0.527, 0.539]; [3, 30]	[0.495, 0.629]; [A3, A1]	[0.502, 0.542]; [E^R_{FD}, E_{GG}]
8	[0.748, 0.761]; [10, 3]	[0.544, 0.844]; [A3, A1]	[0.625, 0.817]; [E_{FD}, E^R_{GG}]
9	[0.457, 0.472]; [3, 5]	[0.435, 0.564]; [A2, A1]	[0.445, 0.475]; [E^R_{FD}, E_{GG}]
10	[0.685, 0.756]; [20, 3]	[0.354, 0.890]; [A3, A1]	[0.686, 0.746]; [$E_{GD} = E_{FD} = E_{LK}$, E^R_{FD}]

TABLE IV
OPERATIONS WITH THE BEST ACCURACY VALUES

data	Reference metrics			Best tested $A(E)$		
set	Chebyshev	Euclidean	Manhattan $\simeq$ ($A1(E_{LK})$)	1st best	2nd best	3rd best
1	0.996	0.996	0.996	**0.997** $A2(E_{LK})p = 0.5$	0.996 $A2(E^R_{LK})p = 2, p = 3$	0.996 $A1(E^R_{LK})p = 0.5$ $A2(E^R_{LK})p = 0.5, p = 4$
2	0.834	0.853	**0.859**	0.858 $A1(E^R_{LK})$	0.852 $A1(E_{GG})$	0.846 $A1(E_{GD})$
3	0.892	0.888	0.89	**0.898** $A1(E_{GG})$	0.896 $A1(E^R_{LK})$, $A2(E_{GG})p = 2$	0.894 $A1(E^R_{FD})$
4	0.616	0.714	0.718	**0.719** $A1(E^R_{GD})$	0.717 $A1(E^R_{LK})$	0.715 $A1(E_{GG})$
5	**0.913**	0.910	0.901	0.901 $A2(E_{LK})$ p=0.5 $A2(E^R_{LK})p = 0.5$ $A2(E_{GG})p = 2, p = 4$	0.9 $A2(E^R_{LK})p = 4$	0.898$A1(E^R_{LK})$ $A2(E_{LK})p = 2$ $A2(E_{GG})p = 3$
6	0.615	0.624	0.639	**0.640** $A1(E^R_{LK})$	0.618 $A1(E_{GG})$	0.613 $A4(E^R_{LK})$
7	0.629	0.661	**0.670**	**0.670** $A1(E^R_{LK})$	0.659 $A1(E_{GG})$	0.615 $A1(E_{GD})$
8	0.871	0.887	0.894	**0.933** $A2(E_{GG})p = 0.5$	0.929 $A1(E_{GG})$	0.920 $A2(E_{GG})p = 2$
9	0.581	**0.596**	0.586	0.585 $A1(E^R_{LK})$	0.578 $A1(E_{GG})$	0.552 $A1(E^R_{GD})$
10	0.506	0.89	0.89	**0.892** $A1(E_{FD})$,$A1(E^R_{FD})$, $A2(E_{FD})$,$A2(E^R_{FD})$ all p	0.89 all remaining $A1(E)$, $A1(E^R)$, $A2(E)$, $A2(E^R)$ all p	0.408 all $A4(E)$, $A4(E^R)$

mining. For example examining properties of a sphere induced by fuzzy equivalences aggregations or, as in paper [19], applying aggregations and fuzzy equivalence to clustering problems.

REFERENCES

[1] L.A. Zadeh, Similarity relations and fuzzy orderings, Inf. Sci. 3 (1971) 177–200.

[2] S. Montes, J. Jiménez, P. Gil, -Partitions and -equivalences, Mathw. Soft Comput. 5 (1998) 201–211.

[3] J.C. Fodor and M. Roubens, Fuzzy preference modelling and multicriteria decision support, Kluwer, Dordrecht 1994.

[4] U. Bentkowska, J.G. Bazan, W. Rząsa, and L. Zaręba, Application of interval-valued aggregation to optimization problem of k-NN classifiers for missing values case, Inf. Sci. 486 (2019), 434–449.

[5] M. Kordestani, A. Chaibakhsh, M. Saif, SMS - a Security Management System for Steam Turbines Using a Multisensor Array, IEEE Systems Journal, 2020.

[6] T. Calvo, A. Kolesárová, M. Komorniková and R. Mesiar, Aggregation operators: Properties, classes and construction methods. In T. Calvo et al., editors, *Aggregation Operators*, pages 3–104, Physica-Verlag, Heildelberg, 2002.

[7] E.P. Klement, R. Mesiar, E. Pap, Triangular norms, Kluwer Academic Publishers, Dordrecht, 2000.

[8] U. Bentkowska, Interval-valued methods in Classifications and Decisions, Studies in Fuzziness and Soft Computing, vol. 378, Springer 2020.

[9] J. Drewniak and A. Król, A survey of weak connectives and the preservation of their properties by aggregations, Fuzzy Sets Syst. 161 (2010), 202–215.

[10] M. Baczyński, B. Jayaram, Fuzzy implications, Springer, Berlin, 2008.

[11] U. Bentkowska and A. Król, Fuzzy α-C-equivalences, Fuzzy Sets Syst. 360 (2019), 49–64.

[12] H. T. Nguyen, E. Walker, A first course in fuzzy logic, CRC Press, Boca Raton, 1996.

[13] A. Król, Fuzzy (C,I)-equivalences, in: Baczyński et al. (eds.), Proceedings of the 8th International Summer School on Aggregation Operators, University of Silesia, Katowice, Poland, pp. 157–161 (2015)

[14] U. Bentkowska and A. Król, Dependencies between some types of fuzzy equivalences, in: J. Medina et al. (eds.), IPMU 2018, CCIS 853 Springer International Publishing AG 2018, part of Springer Nature, 661–672.

[15] U. Bentkowska, B. Pękala, and W. Rząsa, The stability of local properties of fuzzy relations under ordinal equivalence, Inf. Sci. 491 (2019), 265–278.

[16] U. Bentkowska and A. Król, Aggregation of fuzzy α-C-equivalencess, in: Proceedings of the International Joint Conference IFSA-EUSFLAT 2015, J.M. Alonso Atlantis Press 2015, 1310–1317.

[17] UC Irvine Machine Learning Repository, https://archive.ics.uci.edu/ml/datasets.php .

[18] Dell Statistica (data analysis software system) version 13. software.dell.com .

[19] E. Armengol, D. Boixader, A. García-Cerdana, and J. Recasens, T-generable indistinguishability operators and their use for feature selection and classification, Fuzzy Sets Syst. 360 (2019), 33–48.