Attribute Selection for Sets of Data Expressed by Intuitionistic Fuzzy Sets

1st Eulalia Szmidt
Systems Research Institute
Polish Academy of Sciences
Newelska 6, 01-447 Warsaw, Poland
szmidt@ibspan.waw.pl

2nd Janusz Kacprzyk
Systems Research Institute
Polish Academy of Sciences
Newelska 6, 01-447 Warsaw, Poland
kacprzyk@ibspan.waw.pl

3rd Paweł Bujnowski
Systems Research Institute
Polish Academy of Sciences
Newelska 6, 01-447 Warsaw, Poland
pbujno@ibspan.waw.pl

***Abstract*—In this paper we continue our previous considerations on attribute selection while a data set is expressed via an Atanassov's intuitionistic fuzzy set (IFS). We examine a simple, yet it seems, a powerful algorithm making use of the three term attribute description (membership values, non-membership values, and hesitation margins). We test the algorithm on the examples being well known benchmarks data from UCI Machine Learning Repository. The results are compared with one of the most used methods of model reduction, namely, the Principal Component Analysis.**

***Index Terms*—feature selection, intuitionistic fuzzy sets, three term description of the attributes**

I. Introduction

Problem of model dimensionality reduction is still a "hot" one as the possible approaches have their pros and cons but there is not one "the best method". Model reduction can be done by two possible approaches. The first method is known as feature (attribute) extraction when the dimensionality is reduced by using a combination of features (attributes), forming some new ones, and this can be difficult for a user as the method is not transparent and may generate difficulties with model interpretation. The second method is known as feature (attributes) selection and boils down to the selection of the most relevant features. Here we will examine attribute selection for the sets of data expressed by the intuitionistic fuzzy sets (IFSs, for short).

The intuitionistic fuzzy sets (Atanassov [1], [2], [3]) are a very convenient tool for modeling the systems in the presence of a lack of knowledge which is omnipresent practically in (almost) all decision making problems and the same time difficult to foresee. IFSs which are an extension of the fuzzy sets (Zadeh [35]), have an inherent possibility to take into account a lack of knowledge by using the so-called hesitation margin or intuitionistic fuzzy index (cf. Szmidt [15], Szmidt and Kacprzyk [27], [31], [25], [18], [19], Szmidt, Kacprzyk, Bujnowski [29]).

However, as other types of models the IFS models can be described by too many variables to efficiently perform simulations. So, here again, the reduction of dimensionality of data is a problem to solve. The well known Principal Component Analysis (PCA) for the IFSs (cf. Szmidt and Kacprzyk [28]), Szmidt [15]) gives correct results but, again, it is quite complicated from the point of view of calculations, and the final result is not transparent enough for some users.

Here we examine a simple method of feature selection for the data sets which are expressed by intuitionistic fuzzy sets (IFSs). Three term representation of IFSs making use of membership values, non-membership values, and hesitation margins, enables transparent, simple and efficient feature selection. The method is easy from the point of view of calculations. More, the considered here approach makes it possible to rank the attributes (not all methods enable it).

We illustrate and test the discussed method on examples being well known benchmark data from UCI Machine Learning Repository (https://archive.ics.uci.edu/ml/datasets).

The results are compared with Principal Component Analysis being a well known method of a reduction dimensionality of a model (cf. Jackson [9], Jolliffe [10], Marida et al. [12]).

II. A Brief Introduction to IFSs

One of the possible generalizations of a fuzzy set in X (Zadeh [35]) given by

$$A^{'} = \{\langle x, \mu_{A^{'}}(x)\rangle | x \in X\} \tag{1}$$

where $\mu_{A^{'}}(x) \in [0,1]$ is the membership function of the fuzzy set $A^{'}$, is an IFS (Atanassov [1], [2], [3]) A is given by

$$A = \{\langle x, \mu_A(x), \nu_A(x)\rangle | x \in X\} \tag{2}$$

where: $\mu_A : X \to [0,1]$ and $\nu_A : X \to [0,1]$ such that

$$0 \leq \mu_A(x) + \nu_A(x) \leq 1 \tag{3}$$

and $\mu_A(x)$, $\nu_A(x) \in [0,1]$ denote a degree of membership and a degree of non-membership of $x \in A$, respectively. (See Szmidt and Baldwin [16] for assigning memberships and non-memberships for IFSs from data.)

Obviously, each fuzzy set may be represented by the following IFS:
$A = \{\langle x, \mu_{A^{'}}(x), 1 - \mu_{A^{'}}(x)\rangle | x \in X\}$.

An additional concept for each IFS in X, that is not only an obvious result of (2) and (3) but which is also relevant for applications, we will call (Atanassov [2])

$$\pi_A(x) = 1 - \mu_A(x) - \nu_A(x) \tag{4}$$

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

a *hesitation margin* of $x \in A$ which expresses a lack of knowledge of whether x belongs to A or not (cf. Atanassov [2]). It is obvious that $0 \leq \pi_A(x) \leq 1$, for each $x \in X$.

The hesitation margin turns out to be important while considering the distances (Szmidt and Kacprzyk [17], [18], [22], entropy (Szmidt and Kacprzyk [19], [23]), similarity (Szmidt and Kacprzyk [24]) for the IFSs, etc. i.e., the measures that play a crucial role in virtually all information processing tasks (Szmidt [15]).

The hesitation margin turns out to be relevant for applications – in image processing (cf. Bustince et al. [7]), the classification of imbalanced and overlapping classes (cf. Szmidt and Kukier [32], [33], [34]), the classification applying intuitionistic fuzzy trees (cf. Bujnowski [6]), group decision making (e.g., [4]), genetic algorithms [13], negotiations, voting and other situations (cf. Szmidt and Kacprzyk papers).

III. Three term representation of the IFSs as a foundation for attribute selection

In our previous papers we have shown that three term representation of the IFSs, i.e., taking into account membership values μ, non-membership values ν and hesitation margins π is very useful especially from practical points of view (cf. Szmidt [15], Szmidt and Kacprzyk [30], [18], [19], [20], [21], [25], [27], [31]). In this paper we also make use of the same form of IFSs representation.

In [16] we have presented an algorithm of how to derive IFS parameters of a model from relative frequency distributions (histograms). In this paper we assume that the parameters are known.

In the intuitionistic fuzzy model considered in this paper the attributes are described by the above mentioned three terms. Having in mind the interpretation of the three terms we can try to point out the most relevant attributes. As the values of each attribute A_k, $k = 1, \ldots, K$ for different instances are different, an attribute can be described by average values of memberships (5), non-memberships (6), and hesitation margins (7), i.e.:

$$\overline{\mu}_{A_k} = \frac{1}{n} \sum_{i=1}^{n} \mu_{A_k}(x_i) \tag{5}$$

$$\overline{\nu}_{A_k} = \frac{1}{n} \sum_{i=1}^{n} \nu_{A_k}(x_i) \tag{6}$$

$$\overline{\pi}_{A_k} = \frac{1}{n} \sum_{i=1}^{n} \pi_{A_k}(x_i) \tag{7}$$

where n is a number of instances.

Having all the attributes described by (5)– (7) it is possible to point out the most discriminative attributes. A specific intuitionistic fuzzy attribute A_k is most discriminative if its average intuitionistic fuzzy index (7) is as small as possible, and the difference between average membership value and average non-membership value $|\overline{\mu}_{A_k} - \overline{\nu}_{A_k}|$ is as big as possible. The simplest function which makes it possible to find out the most relevant attributes, i.e., the one fulfilling conditions for π and $|\overline{\mu}_{A_k} - \overline{\nu}_{A_k}|$ is:

$$f(A_k) = [(1 - \overline{\pi}_{A_k})(|\overline{\mu}_{A_k} - \overline{\nu}_{A_k}|)] \tag{8}$$

Function $f(A_k)$ (8) which is in fact $f(A_{\overline{\mu}_k, \overline{\nu}_k, \overline{\pi}_k})$ has the following properties

1) $0 \leq f(A_k) \leq 1$.
2) $f(A_k) = (f(A_k)^C)$
3) If a value of $|\overline{\mu_k} - \overline{\nu_k}|$ is fixed, $f(A_k)$ increases while π decreases.
4) If a value of π is fixed, $f(A_k)$ behaves dually to a very simple sort of entropy measure $|\overline{\mu_k} - \overline{\nu_k}|$ (i.e., as $1 - (|\overline{\mu_k} - \overline{\nu_k}|)$).

The shape of (8), and its contour plot are in Figure 1.

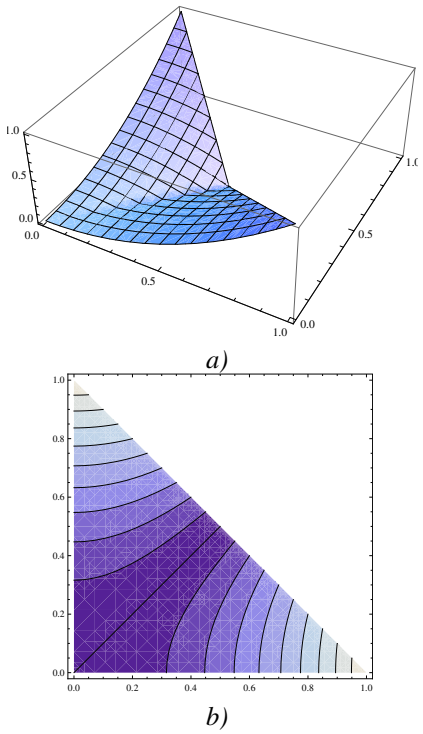

Fig. 1. Function (8): a)- shape; b)- contourplot

It is easy to notice that the shape of $|\overline{\mu_k} - \overline{\nu_k}|$ is always the same in spite of π.

Having characteristic of each attribute $f(A_k)$ (8) we find "the best" attribute

$$\arg\max_{A_k} [(1 - \overline{\pi}_{A_k})(|\overline{\mu}_{A_k} - \overline{\nu}_{A_k}|)] \tag{9}$$

where A_k is the k-th attribute , $k = 1, \ldots, K$.
We can rank all K attributes from the most to the least discriminative by repeating (9) $K - 1$ times.

IV. Results

We will test the method (8)–(9) of the selection of the attributes using several benchmark data sets.

First, using the Weka (http://www.cs.waikato.ac.-nz/ml/weka/) we evaluated the accuracy of different classifiers

TABLE I
"CHESS" – FIRST TEN ATTRIBUTES SELECTED BY $f(A_k)$ (8)

	1	2	3	4	5	6	7	8	9	10
Attribute No	10	33	21	18	7	8	35	13	9	6
measure $f(A_k)$	0.049	0.042	0.033	0.023	0.022	0.021	0.017	0.013	0.011	0.01

using all the attributes (without selection) of a tested data set. A simple cross validation method was used with 10 experiments of the 10-fold cross validation. We tested 12 algorithms obtainable from Weka but present here results of the best ones for each data set. Below there are the tested algorithms which (for different benchmarks) turned out the most accurate.

- ***J48*** – implementation of the crisp tree proposed by Quinlan *C4.5* ([14]);
- ***LMT*** (*Logistic Model Tree*) – a hybrid tree with the logistic models at the leaves; ([11]),
- ***Random Forest*** – here consisting of 10 decision trees with nodes generated on the basis of a random set of attributes ([5]);
- ***Multilayer Perceptron*** – neural network;
- ***Logistic*** – logistic regression;
- ***ADA Boost***;
- ***k-nn classifier*** – k-nn classifier for $n = 1$.

Next, using (8)–(9) we evaluate and rank all the attributes for a data set. Again, using the Weka we assess the accuracy of different classifiers using more and more attributes starting from only one attribute (the best one), and adding one by one the next ranked attributes. After assessing the results of the classification we present results for a chosen numbers of the attributes for which the accuracy is still acceptable. We compare the accuracy for the chosen number of the ranked attributes by (8)–(9) with the results obtained for the same number of the attributes by using Principal Component Analysis (PCA) – one of the best known and widely used linear dimension reduction technique Jackson [9], Jolliffe [10], Marida et al. [12] in the sense of mean-square error.

A. Data Set Chess

The first data set we use to verify the procedure (8)–(9) is "Chess". Data set Chess has 36 attributes. The last (37th) attribute is the classification: "win" or "nowin". There are 3196 instances: 1669 (52%) white can win, and 1527 (48%) white cannot win. In Fig 2 there are all the attributes evaluated by (8) and presented in descended order from the best to the worst one. The order of first 10 best attributes and respective values of measure (8) are in Table I.

Next, using the Weka we evaluated the accuracy of different 12 classifiers using all 36 attributes (without selection). A simple cross validation method was applied with 10 experiments of the 10-fold cross validation. The best results of the classification were obtained for the algorithms:

- trees LMT;
- decision tree J48;

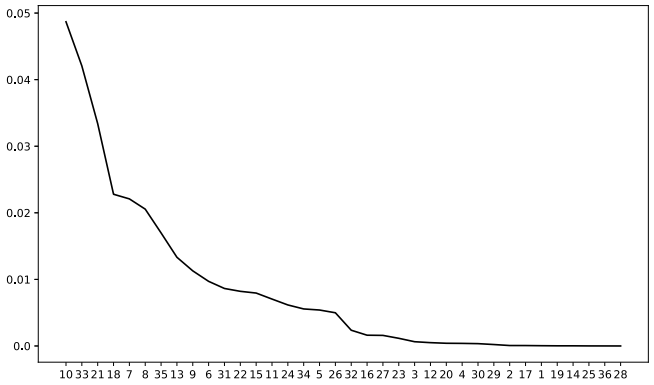

Fig. 2. The values of (8) for all the Chess attributes ranked from the best to the worst

TABLE II
"CHESS" – COMPARISON OF THE CLASSIFICATION ACCURACY BY DIFFERENT CLASSIFIERS WITH ALL 36 ATTRIBUTES

	Classification accuracy ($\bar{x} \pm \sigma$) w %	
Algorithm (no selection)	**accuracy of both classes**	**AUC ROC**
trees LMT	99.64 ± 0.35	1 ± 0.0
decision tree J48	99.44 ± 0.40	0.998 ± 0.003
Random Forest	99.23 ± 0.44	0.999 ± 0.001
Multilayer Perceptron	97.27 ± 0.48	0.999 ± 0.001

- Random Forest;
- Multilayer Perceptron.

TABLE III
"CHESS" – COMPARISON OF THE CLASSIFICATION ACCURACY BY $f(A_k)$ (9), AND PCA REDUCTIONS, WITH 30 ATTRIBUTES

	Classification accuracy ($\bar{x} \pm \sigma$) w %	
Algorithm (30 attributes)	$f(A_k)$ **(9)**	**PCA**
trees LMT	98.14 ± 0.79	96.74 ± 1.02
decision tree J48	97.93 ± 0.80	89.31 ± 1.94
Random Forest	97.82 ± 0.79	96.05 ± 1.22
Multilayer Perceptron	97.54 ± 0.97	98.33 ± 0.57

Besides of the classification accuracy (total proper identification of the instances belonging to the classes considered), we have also paid attention to the area under ROC curve [8]. The results are in Table II. The accuracy by the best algorithms and all the attributes is high (Table II)– 99.64% obtained by a logistic model tree (LMT). Also other algorithms: decision tree J48, Random Forest, and Multilayer Perceptron have similar accuracy, namely, 99,4%, 99.2%, 97.3% respectively. We wanted to see how many attributes is redundant so selecting the attributes we had in mind still the hight accuracy. Taking into account 30 "best" attributes only (Table III) we obtained

TABLE IV
"MUSK" – FIRST 20 ATTRIBUTES SELECTED BY $f(A_k)$ (8)

	1	2	3	4	5	6	7	8	9	10
Attribute No	97	65	132	148	4	17	118	149	138	70
measure $f(A_k)$	0.071	0.070	0.066	0.066	0.064	0.062	0.056	0.055	0.052	0.049
	11	12	13	14	15	16	17	18	19	20
Attribute No	136	98	143	130	59	89	119	96	7	100
measure $f(A_k)$	0.049	0.044	0.044	0.044	0.043	0.043	0.041	0.041	0.040	0.039

accuracy 98.14% for LMT and not less than 97.54% for the fourth algorithm (Multilayer Perceptron). In the same Table II) we have results obtained by PCA. The PCA results are worse for the three first algorithms. In other words the selecting algorithm (8)–(9) meet our expectations for Chess data set.

B. Data Set Musk

Next, we examined data set Musk with 166 attributes and 476 (207 + 269) Instances. A classifier should classify a molecule as "musk" or as "non-musk". Details are here: (https://archive.ics.uci.edu/ml/datasets/Musk

In Fig 3 there are all the attributes evaluated by (8) and presented in descended order from the best to the worst one. The order of the first 20 best attributes and respective values of measure (8) are in Table IV.

In Table V there are results from Weka, i.e., comparison of the classification accuracy by different classifiers while using all 166 attributes. We were testing 12 algorithms but show only the best 4 algorithm results. The best accuracy was obtained by Multilayer Perceptron (94%). The next results were obtained by Random Forest – 89.8%, k-1 classifier – 85.5%, and trees LMT – 85%.

In Table VI there are results of the classification for the same algorithms but with 68 attributes only. We wanted to select as little attributes as possible to obtain not considerably worse results of classification. Pointing out 68 attributes, i.e., 41% of the attributes in the original data set we obtained results with accuracy only 1% worse than for all 166 attributes, namely, using Multilayer Perceptron we received 93%, Random Forest – 88%, k-1 classifier – 84%, and trees LMT – 85%.

In the same Table VI there are results obtained by the same algorithms, the same number of the extracted attributes but via PCA method. The results are worse than for the method (8)–(9). The best algorithm, i.e., Multilayer Perceptron gives results of classification 11% worse while using PCA than while using the method (8)–(9). For other algorithms the difference is not so big (in average 1%) but every time the method (8)–(9) turned out to give better results of the classification than PCA.

C. Data Set QSAR Biodegradation

QSAR Biodegradation is a data set with 41 attributes, and 1055 instances (class proportion: 356:699). Classification is to discriminate ready (356) and not ready (699) biodegradable molecules (https://archive.ics.uci.edu/ml/datasets/QSAR+biodegradation).

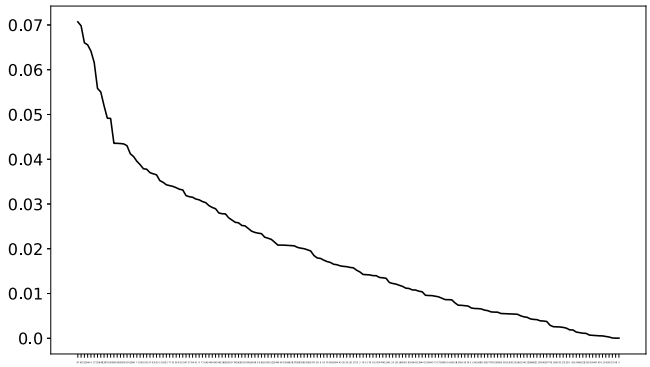

Fig. 3. The values of (8) for all the Musk attributes ranked from the best to the worst

TABLE V
"MUSK" – COMPARISON OF THE CLASSIFICATION ACCURACY BY DIFFERENT CLASSIFIERS WITH ALL 166 ATTRIBUTES

Algorithm (no selection)	**Classification accuracy ($\bar{x} \pm \sigma$) w %**	
	accuracy of both classes	**AUC ROC**
Multilayer Perceptron	94.24 ± 3.77	0.98 ± 0.02
Random Forest	89.84 ± 5.23	0.96 ± 0.03
k-1 classifier	85.47 ± 4.95	0.86 ± 0.05
trees LMT	85.02 ± 6.64	0.91 ± 0.06

TABLE VI
"MUSK" – COMPARISON OF THE CLASSIFICATION ACCURACY BY $f(A_k)$ (9), AND PCA REDUCTIONS, WITH 68 ATTRIBUTES

Algorithm (68 attributes)	**Classification accuracy ($\bar{x} \pm \sigma$) w %**	
	$f(A_k)$ **(9)**	**PCA**
Multilayer Perceptron	93.16 ± 3.89	81.73 ± 5.94
Random Forest	88.03 ± 5.11	85.43 ± 5.22
k-1 classifier	84.20 ± 4.86	83.37 ± 5.41
trees LMT	$85,13 \pm 5.56$	84.83 ± 4.90

In Fig 4 there are all the attributes evaluated by (8) and presented in descended order from the best to the worst one. The order of the first 20 best attributes and respective values of measure (8) are in Table VII.

The results of the classification with all 41 attributes (Table VIII) are worse than for the previous data sets. For the best four algorithm the accuracy is approximately equal to 87%.

The results of the classification with selected 19 attributes by (8)–(9), i.e., with 46% of the original number of the attributes, are in Table IX. For the best algorithm (function SGD) the accuracy is 85% (with all the attributes it was 87%). Accuracy while using Random Forest is 86% (with all the

TABLE VII
"QSAR BIODEGRADATION" – FIRST 20 ATTRIBUTES SELECTED BY $f(A_k)$ (8)

	1	2	3	4	5	6	7	8	9	10
Attribute No	7	33	39	38	34	3	8	1	15	25
measure $f(A_k)$	0.056	0.043	0.042	0.041	0.033	0.029	0.028	0.019	0.019	0.018
	11	12	13	14	15	16	17	18	19	20
Attribute No	14	30	27	2	31	13	36	11	22	10
measure $f(A_k)$	0.017	0.017	0.016	0.015	0.010	0.010	0.010	0.01	0.010	0.009

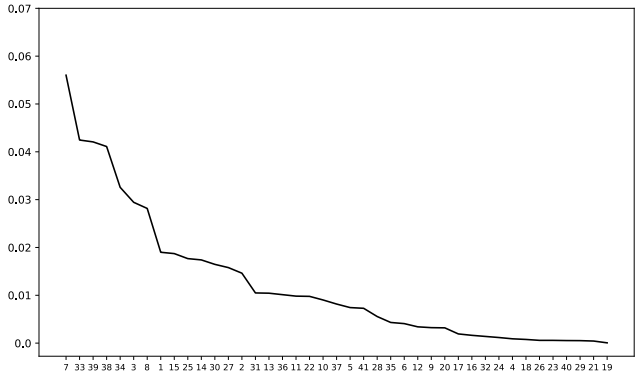

Fig. 4. The values of (8) for all the QSAR Biodegradation attributes ranked from the best to the worst

TABLE VIII
"QSAR BIODEGRADATION" – COMPARISON OF THE CLASSIFICATION ACCURACY BY DIFFERENT CLASSIFIERS WITH ALL 41 ATTRIBUTES

	Classification accuracy ($\bar{x} \pm \sigma$) w %	
Algorithm (no selection)	**accuracy of both classes**	**AUC ROC**
function SGD	87.09 ± 3.38	0.85 ± 0.04
Random Forest	86.88 ± 3.26	0.93 ± 0.03
function Logistic	86.70 ± 3.36	0.92 ± 0.03
trees LMT	86.26 ± 3.35	0.92 ± 0.03

attributes it was 87%). Function Logistic gives accuracy 84% instead of 87% with all the attributes. Trees LMT give 2% worse accuracy (84% instead of 86% with all the attributes). In general the results of the classification are 1%–2% worse with 19 attributes in comparison to the results of classification obtained for all 41 attributes.

In Table IX there are also results of classification with 19 attributes but extracted by PCA. For the best algorithm (function SGD) the method (8)–(9) gives better result than PCA (85% instead of 84%). Random Forest gives almost the same accuracy (approximately 86%). The same situation is while using trees LMT – both (8)–(9) and PCA give accuracy 84%.

However, it is worth noticing that the method (8)–(9) is transparent (selection of the attributes) whereas extraction of the attributes by PCA is like a black box for a user.

D. Data Set Ionosphere

Ionosphere is a data set with 34 attributes. The 35th attribute is either "good" or "bad" according to radar returns from the ionosphere. Number of Instances 351. Class

TABLE IX
"QSAR BIODEGRADATION" – COMPARISON OF THE CLASSIFICATION ACCURACY BY $f(A_k)$ (9), AND PCA REDUCTIONS, WITH 19 ATTRIBUTES

	Classification accuracy ($\bar{x} \pm \sigma$) w %	
Algorithm (19 attributes)	$f(A_k)$ **(9)**	**PCA**
function SGD	84.90 ± 3.48	83.86 ± 3.46
Random Forest	85.60 ± 3.57	86.40 ± 3.12
function Logistic	83.90 ± 3.38	84.74 ± 3.54
trees LMT	$84,18 \pm 3.41$	84.38 ± 3.64

proportion: 225:126.
See: https://archive.ics.uci.edu/ml/datasets/ionosphere for more details.

In Fig 5 there are all the attributes evaluated by (8) and presented in descended order from the best to the worst one. The order of the first 20 best attributes and respective values of measure (8) are in Table X.

In Table XI there is classification accuracy by different classifiers with all 34 attributes. Random Forest gives accuracy 94%, trees LMT – 93%, Multilayer Perceptron – 91%, and ADA Boost – 91%.

In Table XII there are results of classification with 6 attributes, i.e., with only 18% of the original number of the attributes. The substantial reduction of the number of the attributes by (8)–(9) influences the classification accuracy but less than we could expect. Random Forest accuracy is 93% (we had 94% with all the attributes). Trees LMT accuracy is 92% instead of 93%. Multilayer Perceptron gives 90% accuracy (with all the attributes – 91%). ADA Boost classifies instances with 90% accuracy (with all the attributes it was 91%). The results with the substantial reduction of the attributes are only 1% worse.

In Table XII there are also results obtained by PCA, For Random Forest the PCA results are worse (91%) than obtained from (8)–(9) – 93%. Trees LMT and PCA produce 89% of accuracy whereas from (8)–(9) it was 92%. Only in a case of Multilayer Perceptron and PCA we obtain 91% of accuracy whereas for (8)–(9) we have 90%. For ADA Boost again PCA is worse with 87% of accuracy whereas (8)–(9) is better with 90% of accuracy.

V. CONCLUSIONS

We have discussed a method for feature selection for the sets of data represented by the intuitionistic fuzzy sets (IFS). We used the three term model of the IFSs i.e., taking into account the degree of membership, non-membership and hesitation

TABLE X
"IONOSPHERE" – FIRST 20 ATTRIBUTES SELECTED BY $f(A_k)$ (8)

	1	2	3	4	5	6	7	8	9	10
Attribute No	2	32	26	27	3	4	20	30	19	6
measure $f(A_k)$	0.068	0.065	0.061	0.060	0.053	0.044	0.039	0.032	0.029	0.028
	11	12	13	14	15	16	17	18	19	20
Attribute No	25	17	13	2	14	8	23	5	7	9
measure $f(A_k)$	0.025	0.025	0.024	0.024	0.023	0.022	0.022	0.022	0.020	0.018

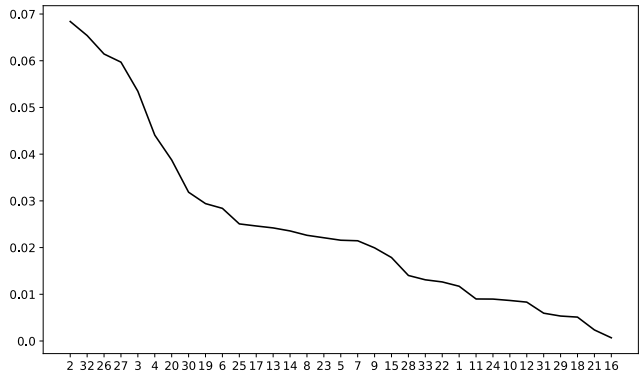

Fig. 5. The values of (8) for all the Ionosphere attributes ranked from the best to the worst

TABLE XI
"IONOSPHERE" – COMPARISON OF THE CLASSIFICATION ACCURACY BY DIFFERENT CLASSIFIERS WITH ALL 34 ATTRIBUTES

	Classification accuracy ($\bar{x} \pm \sigma$) w %	
Algorithm (no selection)	**accuracy of both classes**	**AUC ROC**
Random Forest	93.59 ± 4.02	0.92 ± 0.02
Trees LMT	92.91 ± 4.10	0.84 ± 0.06
Multilayer Perceptron	91.17 ± 4.87	0.93 ± 0.06
ADA Boost	90.89 ± 4.64	0.95 ± 0.05

margin. This description makes it possible to formulate a very natural and understandable function for the evaluation of the attributes being a foundation of the selection process. In this paper it was shown that the considered function (8) is useful from the point of view of classification problems. However, it is worth stressing that the function can be useful in other domains while pointing out the most important attributes and/or ordering them is of paramount importance. Broadly understood decision making is an area where the presented method can be applied. The method is transparent, simple from the point of view of calculations and gives promising results.

REFERENCES

[1] Atanassov K. (1983), Intuitionistic Fuzzy Sets. VII ITKR Session. Sofia (Centr. Sci.-Techn. Libr. of Bulg. Acad. of Sci., 1697/84) (in Bulgarian).

[2] Atanassov K. T. (1999), Intuitionistic Fuzzy Sets: Theory and Applications. Springer-Verlag.

[3] Atanassov K.T. (2012) On Intuitionistic Fuzzy Sets Theory. Springer-Verlag.

[4] Atanassova V. (2004) Strategies for Decision Making in the Conditions of Intuitionistic Fuzziness. Int. Conf. 8th Fuzzy Days, Dortmund, Germany, 263–269.

[5] Breiman L. (2001): *Random Forests.* Machine Learning 45 (1): 50–32.

TABLE XII
"IONOSPHERE" – COMPARISON OF THE CLASSIFICATION ACCURACY BY $f(A_k)$ (9), AND PCA REDUCTIONS, WITH 6 ATTRIBUTES

	Classification accuracy ($\bar{x} \pm \sigma$) w %	
Algorithm (6 attributes)	$f(A_k)$ **(9)**	**PCA**
Random Forest	92.68 ± 4.02	91.25 ± 3.99
Trees LMT	91.58 ± 4.59	88.7 ± 4.86
Multilayer Perceptron	89.46 ± 4.61	90.95 ± 4.41
ADA Boost	$90,18 \pm 4.75$	87.13 ± 5.30

[6] Bujnowski P., Szmidt E., Kacprzyk J. (2014): Intuitionistic Fuzzy Decision Trees - a new Approach. In: Rutkowski L., Korytkowski M., Scherer R., Tadeusiewicz R., Zadeh L., Zurada J. (Eds.): Artificial Intelligence and Soft Computing, Part I. Springer, Switzerland, 181-192.

[7] Bustince H., Mohedano V., Barrenechea E., Pagola M. (2006) An algorithm for calculating the threshold of an image representing uncertainty through A-IFSs. IPMU'2006, 2383–2390.

[8] Hand D. J., Till R. J. (2001): *A simple generalization of the area under the ROC curve for multiple class classification problems.* Machine Learning, 45, 171–186.

[9] Jackson J.E. (1991) A User's Guide to Principal Components. New York: John Wiley and Sons.

[10] Jolliffe I.T. (1986) Principal Component Analysis. Springer-Verlag, 1986.

[11] Landwehr N., Hall M., Frank E. (2005): *Logistic Model Trees.* Machine Learning. 95(1-2), 161–205.

[12] Mardia K.V., Kent J.T., Bibby J.M. (1995) Multivariate Analysis. Probability and Mathematical Statistics. Academic Press, 1995.

[13] Roeva O. and Michalikova A. (2013) Generalized net model of intuitionistic fuzzy logic control of genetic algorithm parameters. In: Notes on Intuitionistic Fuzzy Sets. Academic Publishing House, Sofia, Bulgaria. Vol. 19, No. 2, 71–76. ISSN 1310-4926.

[14] Quinlan J. R. (1993): *C4.5: Programs for Machine Learning.* San Mateo, Morgan Kaufman Publishers, Inc.

[15] Szmidt E. (2014): Distances and Similarities in Intuitionistic Fuzzy Sets. Springer.

[16] Szmidt E. and Baldwin J. (2006): *Intuitionistic Fuzzy Set Functions, Mass Assignment Theory, Possibility Theory and Histograms.* 2006 IEEE World Congress on Computational Intelligence, 237–243.

[17] Szmidt E. and Kacprzyk J. (1997) On measuring distances between intuitionistic fuzzy sets, Notes on IFS, 3(4), 1–13.

[18] Szmidt E. and Kacprzyk J. (2000) Distances between intuitionistic fuzzy sets, Fuzzy Sets and Systems, 114(3), 505–518.

[19] , Szmidt E., Kacprzyk J. (2001) Entropy for intuitionistic fuzzy sets. Fuzzy Sets and Systems, 118 (3), 467–477.

[20] Szmidt E. and Kacprzyk J. (2004) A Concept of Similarity for Intuitionistic Fuzzy Sets and its use in Group Decision Making. 2004 IEEE Conf. on Fuzzy Systems, Budapest, 1129–1134.

[21] Szmidt E. and Kacprzyk J. (2005) New Measures of Entropy for Intuitionistic Fuzzy Sets. Notes on Intuitionistic Fuzzy Sets, 11 (2), 12–20.

[22] Szmidt E. and Kacprzyk J. (2006) Distances Between Intuitionistic Fuzzy Sets: Straightforward Approaches may not work. IEEE IS'06, 716–721.

[23] Szmidt E. and Kacprzyk J. (2007). Some problems with entropy measures for the Atanassov intuitionistic fuzzy sets. Applications of Fuzzy Sets Theory. LNAI 4578, 291–297. Springer-Verlag.

[24] Szmidt E. and Kacprzyk J. (2007a). A New Similarity Measure for Intuitionistic Fuzzy Sets: Straightforward Approaches may not work. 2007 IEEE Conf. on Fuzzy Systems, 481–486.

[25] Szmidt E. and Kacprzyk J.: Ranking of Intuitionistic Fuzzy Alternatives in a Multi-criteria Decision Making Problem. In: Proceedings of the conference: NAFIPS 2009, Cincinnati, USA, June 14- 17, 2009, IEEE, ISBN: 978-1-4244-4577-6.

[26] Szmidt E. and Kacprzyk J. (2010) Dealing with typical values via Atanassov's intuitionistic fuzzy sets. Int. J. of General Systems, 39, (5), 489–506.

[27] Szmidt E., Kacprzyk J. (2011) Intuitionistic fuzzy sets – Two and three term representations in the context of a Hausdorff distance. Acta Universitatis Matthiae Belii, Series Mathematics, 19, 2011, 53–62, available at https://actamath.savbb.sk/pdf/acta1908.pdf

[28] Szmidt E., Kacprzyk J. (2012) A New Approach to Principal Component Analysis for Intuitionistic Fuzzy Data Sets. S. Greco et al. (Eds.): IPMU 2012, Part II, CCIS 298, Springer-Verlag Berlin Heidelberg, 529–538.

[29] Szmidt E., Kacprzyk J. and Bujnowski P. (2014) How to measure the amount of knowledge conveyed by Atanassov's intuitionistic fuzzy sets. Information Sciences 257 (2014) 276 – 285.

[30] Szmidt E., Janusz Kacprzyk J. (2015) Two and three term representations of intuitionistic fuzzy sets: Some conceptual and analytic aspects. IEEE International Conference on Fuzzy Systems FUZZ-IEEE 2015: 1-8.

[31] Szmidt E. and Kacprzyk J. (2017) A Perspective on Differences Between Atanassov's Intuitionistic Fuzzy Sets and Interval-Valued Fuzzy Sets. In: Fuzzy Sets, Rough Sets, Multisets and Clustering (Eds.: Vicenc Torra, Anders Dahlbom, Yasuo Narukawa). Series: Studies in Computational Intelligence, Volume 671, 221 – 237, Springer 2017.

[32] Szmidt E. and Kukier M. (2006): Classification of Imbalanced and Overlapping Classes using Intuitionistic Fuzzy Sets. IEEE IS'06, London, 722–727.

[33] Szmidt E. and Kukier M. (2008) A New Approach to Classification of Imbalanced Classes via Atanassov's Intuitionistic Fuzzy Sets. In: Hsiao-Fan Wang (Ed.): Intelligent Data Analysis : Developing New Methodologies Through Pattern Discovery and Recovery. Idea Group, 85–101.

[34] Szmidt E., Kukier M. (2008) Atanassov's intuitionistic fuzzy sets in classification of imbalanced and overlapping classes. In: Panagiotis Chountas, Ilias Petrounias, Janusz Kacprzyk (Eds.): Intelligent Techniques and Tools for Novel System Architectures. Springer, Berlin Heidelberg 2008, 455–471.

[35] Zadeh L.A. (1965) Fuzzy sets. Information and Control, 8, 338–353.