Atlantis Studies in Uncertainty Modelling, volume 1

11th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT 2019)

A New Method for Attributes Selection in Intuitionistic Fuzzy Models

Eulalia Szmidt[a,b], **Janusz Kacprzyk**[a,b] **and Paweł Bujnowski**[a]
[a]Systems Research Institute Polish Academy of Sciences
ul. Newelska 6, 01-447 Warsaw, Poland
[b]Warsaw School of Information Technology, ul. Newelska 6, 01-447 Warsaw, Poland
e-mails: {szmidt, kacprzyk, pbujno}@ibspan.waw.pl

Abstract

We present a new method for attribute selection for models with attribute values expressed via Atanassov's intuitionistic fuzzy sets (IFSs, for short). The three term representation of the IFSs makes it possible to construct a transparent and justified function that makes it possible to select attributes for broadly understood decision making, e.g., for classification. We demonstrate the proposed method using a well known, simple and illustrative example from the literature, and then we show the results for the SPECT Heart data, a real medical preprocessed data set [9]. The data set is quite complicated and difficult to analyze. The results obtained are compared with other methods, and found to be promising.

Keywords: Atanassov's intuitionistic fuzzy sets, Three term representation of A-IFSs, Selection of attributes.

The problem of model dimensionality reduction has been investigated for a long time and, in spite of new approaches constantly proposed (with their pros and cons), the process is still far from a definite solution and there is no overall "best method". We have two possible approaches to the model dimensionality reduction. First, the so called feature (attribute) extraction when the dimensionality is reduced by using a combination of features (attributes) which may result in difficulties with model interpretation. Second, the so called feature (attribute) selection when only the most relevant features are selected and used. Here we will consider the latter, the attribute selection for data sets which are expressed by Atanassov's intuitionistic fuzzy sets (IFSs for short).

The intuitionistic fuzzy sets (Atanassov [1], [2], [3]), are a very convenient tool for the modeling of systems in the presence of some types of imperfect knowledge which is crucial for decision making and at the same time difficult to foresee. The IFSs, being an extension of the fuzzy sets (Zadeh [39]), can make it possible to take into account some additional aspects of imperfect knowledge by making use of, first, the membership and non-membership degrees (which can conveniently model judgments and testimonies in favor and against, or pros and cons), and second, the so-called hesitation margin or intuitionistic fuzzy index (which can represent a lack of knowledge).

However, for non-trivial practical cases, the IFS models can be described by too many variables to efficiently perform simulations. So, here again, we face the well known problem of the reduction of dimensionality of data, or feature selection in our context. The well known Principal Component Analysis (PCA) for the IFSs (cf. Szmidt and Kacprzyk [32]), Szmidt [23]) 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 propose a novel and simple method of feature selection for the data sets which are expressed by intuitionistic fuzzy sets (IFSs). The three term representation of the IFSs makes possible a simple and efficient feature selection process. The method is transparent and easy from the point of view of calculations. Moreover, the proposed approach makes it possible to rank the attributes (not all methods can do this).

We first test the proposed method using a simple yet illustrative benchmark example well known from the literature. Then, we show calculations for a larger real medical example. The results are compared with those obtained by other methods of dimensionality reduction.

1 A brief introduction to the IFSs

One of the possible generalizations of a fuzzy set in X (Zadeh [39]) given by

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

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

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

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

$$0 \le \mu_A(x) + \nu_A(x) \le 1 \quad (3)$$

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 [24] for deriving memberships and non-memberships for A-IFSs from data.)

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])

$$\pi_A(x) = 1 - \mu_A(x) - \nu_A(x) \quad (4)$$

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 $0 \le \pi_A(x) \le 1$, for each $x \in X$.

The hesitation margin turns out to be important while considering the distances (Szmidt and Kacprzyk [25], [26], [28], entropy (Szmidt and Kacprzyk [27], [29]), similarity (Szmidt and Kacprzyk [30]) for the IFSs, etc. i.e., the measures that play a crucial role in virtually all information processing tasks (Szmidt [23]).

The hesitation margin turns out to be relevant for applications – in image processing (cf. Bustince et al. [8]), the classification of imbalanced and overlapping classes (cf. Szmidt and Kukier [34], [35], [36]), the classification applying intuitionistic fuzzy trees (cf. Bujnowski [7]), group decision making (e.g., [4]), genetic algorithms [20], negotiations, voting and other situations (cf. Szmidt and Kacprzyk papers).

In our further considerations we will use the notion of a complement set A^C

$$A^C = \{\langle x, \nu_A(x), \mu_A(x), \pi_A(x)\rangle | x \in X\} \quad (5)$$

Below, because of space limitation we present only necessary materials directing a reader to respective literature.

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

In [24] we have presented an algorithm of how to derive IFS parameters of a model from relative frequency distributions (histograms). To justify the (automatic) method, we have shown some similarities/parallels between the intuitionistic fuzzy set theory and mass assignment theory – a well known tool for dealing with both the probabilistic and fuzzy uncertainties (the proof is in Baldwin at al. [5]). Next step of our approach was to recall a semantics for membership functions – the interpretation having its roots in the possibility theory. Finally, in [24] we have proposed the automatic algorithm assigning all three terms (memberships, non-memberships and hesitation margins) describing the intuitionistic fuzzy sets.

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 (6), non-memberships (7), and hesitancy margins (8), i.e.:

$$\overline{\mu}_{A_k} = \frac{1}{n}\sum_{i=1}^{n} \mu_{A_k}(x_i) \quad (6)$$

$$\overline{\nu}_{A_k} = \frac{1}{n}\sum_{i=1}^{n} \nu_{A_k}(x_i) \quad (7)$$

$$\overline{\pi}_{A_k} = \frac{1}{n}\sum_{i=1}^{n} \pi_{A_k}(x_i) \quad (8)$$

where n is a number of instances.

The most relevant attributes should be most discriminative. For a specific intuitionistic fuzzy attribute A_k it means that its average intuitionistic fuzzy index (8) should be as small as possible, and the difference between average membership value and average non-membership value $|\overline{\mu}_{A_k} - \overline{\nu}_{A_k}|$ should be as big as possible. The simplest function which fulfills such conditions for A_k is:

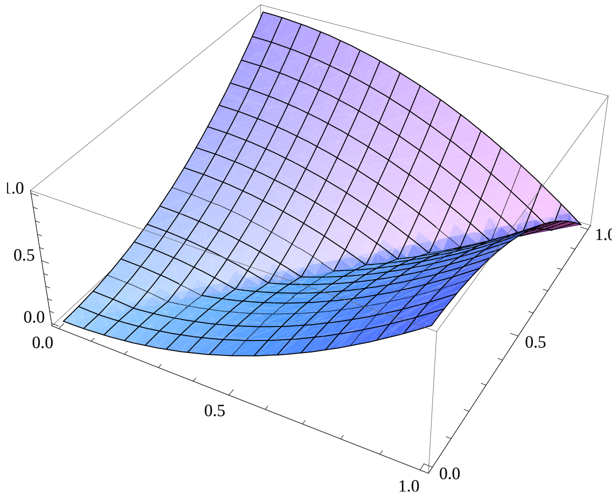
$$f(A_k) = [(1 - \overline{\pi}_{A_k})(|\overline{\mu}_{A_k} - \overline{\nu}_{A_k}|)] \quad (9)$$

Having in mind that function $f(A_k)$ is in fact $f(A_{\overline{\mu}_k, \overline{\nu}_k, \overline{\pi}_k})$, the properties of (9) are:

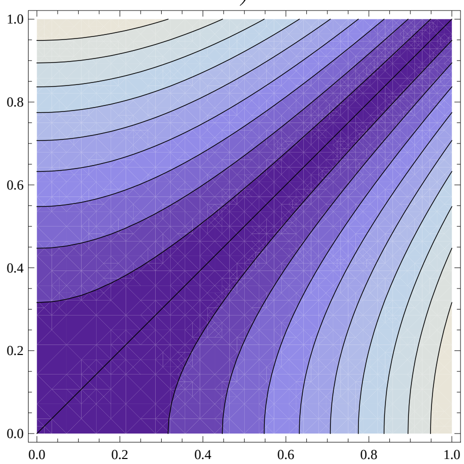
1. $0 \le f(A_k) \le 1$.
2. $f(A_k) = (f(A_k)^C)$
3. For a fixed value of $|\overline{\mu_k} - \overline{\nu_k}|$, $f(A_k)$ increases while π decreases.

4. For a fixed value of π, $f(A_k)$ behaves dually to a very simple sort of entropy measure $|\overline{\mu_k} - \overline{\nu_k}|$ (i.e., as $1 - (|\overline{\mu_k} - \overline{\nu_k}|)$).

In Figure 1 we can see the shape of (9), and its contour plot.



a)



b)

Figure 1: Function (9): a)- shape; b)- contourplot

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

From (9) we find "the best" attribute

$$\arg\max_{A_k}[(1 - \overline{\pi}_{A_k})(|\overline{\mu}_{A_k} - \overline{\nu}_{A_k}|)] \tag{10}$$

where A_k is the k-th attribute , $k = 1, \ldots, K$. Repeating (10) $K - 1$ times we can order all K attributes: from the most to the least discriminative.

3 Results

As an illustration how the proposed method works, we will recall a well known problem formulated by Quinlan [21] but expressed in terms of the IFSs. Quinlan's example, the so-called "Saturday Morning" example, considers the classification with nominal data. This example is small enough and illustrative, yet is a challenge to many classification and machine learning methods. The main idea of solving the example by Quinlan was to select the best attributes (variables) to split the training set (Quinlan used a so-called *Information Gain* which was a dual measure to Shannon's entropy).

No.	Attributes				Class
	Outlook	Temp	Humidity	Windy	
1	sunny	hot	high	false	N
2	sunny	hot	high	true	N
3	overcast	hot	high	false	P
4	rain	mild	high	false	P
5	rain	cool	normal	false	P
6	rain	cool	normal	true	N
7	overcast	cool	normal	true	P
8	sunny	mild	high	false	N
9	sunny	cool	normal	false	P
10	rain	mild	normal	false	P
11	sunny	mild	normal	true	P
12	overcast	mild	high	true	P
13	overcast	hot	normal	false	P
14	rain	mild	high	true	N

Table 1: The "Saturday Morning" data from [21]; N, P, like in [21], stand for a negative, and positive class, respectively

In Quinlan's example [21] (Table 1) we have objects described by attributes. Each attribute represents a feature and takes on discrete, mutually exclusive values. For example, if the objects were "Saturday Mornings" and the classification involved the weather, possible attributes might be [21]:

- **outlook**, with values {sunny, overcast, rain},
- **temperature**, with values {cold, mild, hot},
- **humidity**, with values {high, normal}, and
- **windy**, with values {true, false},

The limitation of space does not let us discuss the method of deriving the IFS counterpart of Quinlan's example (Table 2) in detail (cf. Szmidt and Kacprzyk [31]) and we only present here the final results. Next, making use of the IFS model (Table 2) we compute an average characteristic of the attributes. The results are in Table 3. using the first three attributes as pointed out by our method (Outlook, Humidity, Windy).

Another selection of the attributes obtained via Hellwig's method [12] adapted to IFS data (cf. Szmidt

No.	Attributes				Class
	Outlook	Humidity	Windy	Temperature	
1	(0, 0.33, 0.67)	(0, 0.33, 0.67)	(0.2, 0, 0.8)	(0, 0.33, 0.67)	N
2	(0, 0.33, 0.67)	(0, 0.33, 0.67)	(0, 0.33, 0.67)	(0, 0.33, 0.67)	N
3	(1, 0, 0)	(0, 0.33, 0.67)	(0.2, 0, 0.8)	(0, 0.33, 0.67)	P
4	(0.2, 0.11, 0.69)	(0, 0.33, 0.67)	(0.2, 0, 0.8)	(0, 0, 1)	P
5	(0.2, 0.11, 0.69)	(0.6, 0, 0.4)	(0.2, 0, 0.8)	(0.4, 0.11, 0.49)	P
6	(0.2, 0.11, 0.69)	(0.6, 0, 0.4)	(0, 0.33, 0.67)	(0.4, 0.11, 0.49)	N
7	(1, 0, 0)	(0.6, 0, 0.4)	(0, 0.33, 0.67)	(0.4, 0.11, 0.49)	P
8	(0, 0.33, 0.67)	(0, 0.33, 0.67)	(0.2, 0, 0.8)	(0, 0, 1)	N
9	(0, 0.33, 0.67)	(0.6, 0, 0.4)	(0.2, 0, 0.8)	(0.4, 0.11, 0.49)	P
10	(0.2, 0.11, 0.69)	(0.6, 0, 0.4)	(0.2, 0, 0.8)	(0, 0, 1)	P
11	(0, 0.33, 0.67)	(0.6, 0, 0.4)	(0, 0.33, 0.67)	(0, 0, 1)	P
12	(1, 0, 0)	(0, 0.33, 0.67)	(0, 0.33, 0.67)	(0, 0, 1)	P
13	(1, 0, 0)	(0.6, 0, 0.4)	(0.2, 0, 0.8)	(0, 0.33, 0.67)	P
14	(0.2, 0.11, 0.69)	(0, 0.33, 0.67)	(0, 0.33, 0.67)	(0, 0, 1)	N

Table 2: The “Saturday Morning” data in terms of the IFSs

1	2	3	4	5	6
Attribute	$\overline{\pi}$	$\overline{\mu}$	$\overline{\nu}$	$\lvert\overline{\mu}-\overline{\nu}\rvert$	$(1-\overline{\pi})(\lvert\overline{\mu}-\overline{\nu}\rvert$
Outlook (A_1)	0.48	0.36	0.16	0.2	0.104
Humidity (A_2)	0.53	0.30	0.17	0.13	0.061
Windy (A_3)	0.74	0.11	0.14	0.03	0.008
Temperature (A_4)	0.76	0.11	0.13	0.02	0.005

Table 3: Characterization of the “Saturday Morning” attributes

and Kacprzyk [33]) and verified using the Quinlan's example [33] again gave the same result.

Finally, we compared the new results for the same example but obtained by Principal Component Analysis (PCA) – one of the best known and widely used linear dimension reduction technique Jackson [13], Jolliffe [14], Marida et al. [18] in the sense of mean-square error.

After performing the PCA adapted to the data expressed via the IFSs (cf. Szmidt and Kacprzyk [32]), we have noticed that the first three eigenvalues explain most of variability of the data (85%), and summarize the most important features of the data. It is again a sort of confirmation of our result while using the novel method (10) presented here (PCA as a method of reduction points out a combination of attributes not the initial attributes).

The above example was just for demonstration of the method proposed. As feature selection makes sense for large problems, now we will verify how the method proposed behaves using a real database describing diagnosing of cardiac Single Proton Emission Computed Tomography (SPECT) images [9]. The patients are classified into two categories: normal and abnormal, i.e., with a variety of heart diseases. The database contains 276 instances (originally: images describing 276 patients). The data were preprocessed to extract features summarizing the original SPECT images [9]. In result 22 binary feature patterns (22 attributes for each patient) were derived [9].

First, we used WEKA (http://www.cs.waikato.ac.-nz/ml/weka/) to classify the data with all 22 attributes. Simple cross validation method was used with 10 experiments of 10-fold cross validation. The following algorithms were tested:

- ***J48*** – implementation of the crisp tree proposed by Quinlan *C4.5* ([22]);
- ***LMT*** (*Logistic Model Tree*) – a hybrid tree with the logistic models at the leaves; ([17]),
- ***Random Forest*** – here consisting of 10 decision trees with nodes generated on the basis of a random set of attributes ([6]);
- ***Multilayer Perceptron*** – neural network;
- ***Logistic*** – logistic regression;
- ***Support Vector***;
- ***k-nn classifier*** – k-nn classifier for $n = 1$.

Besides of the classification accuracy (total proper identification of the instances belonging to the classes

Algorithm (no selection)	Classification accuracy $(\bar{x} \pm \sigma)$ w %	
	accuracy of both classes	AUC ROC
C4.5 decision tree (pruned)	81.35 ± 6.16	0.79 ± 0.12
treesLMT	83.95 ± 6.36	0.84 ± 0.10
Random Forest	81.81 ± 6.49	0.79 ± 0.11
Multilayer Perceptron	79.94 ± 6.55	0.80 ± 0.11
LogisticRegression	82.30 ± 6.62	0.83 ± 0.09
Support Vector	83.69 ± 5.70	0.73 ± 0.11
k-nn classifier	81.85 ± 6.13	0.74 ± 0.13

Table 4: "SPECT Heart" – comparison of the results by different classifiers with all 22 attributes

	1	2	3	4	5	6	7	8	9	10
Attribute No	13	8	21	22	16	3	20	10	1	7
measure $f(A_k)$	0.128	0.092	0.086	0.085	0.065	0.061	0.057	0.055	0.048	0.046

Table 5: 'SPECT Heart" – first ten attributes selected by $f(A_k)$ (9)

Algorithm (8 attributes only)	Classification accuracy $(\bar{x} \pm \sigma)$ w %	
	accuracy of both classes	AUC ROC
C4.5 decision tree (pruned)	83.57 ± 5.27	0.80 ± 0.11
treesLMT	84.51 ± 5.20	0.83 ± 0.10
Random Forest	82.97 ± 5.20	0.79 ± 0.10
Multilayer Perceptron	85.00 ± 5.65	0.79 ± 0.11
LogisticRegression	84.85 ± 5.31	0.83 ± 0.10
Support Vector	82.45 ± 5.46	0.70 ± 0.11
k-nn classifier	84.06 ± 5.25	0.78 ± 0.11

Table 6: "SPECT Heart" – comparison of the results by different classifiers with selected 8 attributes

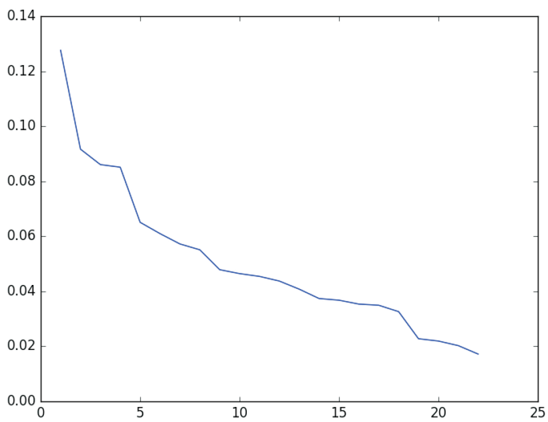


Figure 2: The values of (9) for all the attributes ranked from the best to the worst

considered), we have also compared the area under ROC curve [11]. The results are in Table 4.

Table 4 shows that the accuracy of classification even while using all the attributes is not too high. The highest accuracy while using all the attributes of SPECT Heart was obtained for trees LMT (83.95%). However, detailed analysis of the results shows that 197 instances (out of 212) from the bigger class were classified correctly (15 instances incorrectly) whereas only 26 instances (out of 55) from the smaller class were classified correctly (29 elements from the smaller class, i.e., more than 50% were classified incorrectly). The conclusion is that even while using all the attributes it is difficult to classify the instances of SPECT Heart data, especially the more important instances from the smaller class (ill patients).

Next, we have assessed all the attributes using the proposed function (9). The data [9], as in the previous example, were expressed in terms of IFSs (cf. Szmidt and Baldwin [24]). After performing (6)–(10), we obtained ranking of the attributes in terms of (9). The results for 10 the best attributes (with the highest values of (9)) are in Table 5. In Fig 2 there are all the attributes evaluated by (9) and presented in descended order from the best to the worst one.

In the next part of the experiment we were verifying how many attributes give good classification results. It turned out that 8 first attributes selected by (9), namely: 13, 8, 21, 22, 16, 3, 20 and 10 (Table 5) gives the best results. Both for less and for more attributes the accuracy of the classification was worse. The results of classification with the best 8 attributes, are in Table 6.

It is easy to notice that using only 8 attributes (instead original 22) improves the results of the accuracy for all the tested algorithms. Also the results for ROC curves are very similar (the best for k-nn classifier) and not less stable (with the same or even smaller standard deviations).

The obtained results confirm that the proposed feature selection method can be used to identify and remove irrelevant and redundant attributes from data that do not contribute to the accuracy of a model or even may decrease the accuracy of the model.

Two other approaches of feature selection, so called CLIP3 and CLIP4 algorithms, which were tested using the same data, are presented in [16] and [10]. Their accuracy is 84% and 86.1%, respectively (as compared with cardiologists' diagnoses). CLIP4 [10] discovered that only a subset of 7 out of 22 attributes were strongly relevant, namely: { 21, 18, 17, 16, 10, 8, 7}. All the selected strongly relevant features were discovered to be equally relevant from the point of view of classification.

However, the accuracy obtained by CLIP4, i.e., 86% is also not very satisfactory especially as far as pointing out instances from the smaller class, i.e., ill patients. We have verified that using the subset of the attributes obtained by CLIP 4: 21, 18, 17, 16, 10, 8, 7, the best algorithm i.e., Logistic Regression (among the tested WEKA algorithms), can see properly only 37 ill patients whereas 18 ill patients were classified as normal cases.

In other words, the data base is difficult for standard classification algorithms (Table 4). The same conclusion is confirmed in [19]. Selection of the attributes presented in [19] give substantially higher accuracy (up tp 96%) but the subset of the attributes contains 15 elements (LIP4 pointed out 7 elements only). This way it has been shown [19] that there are redundant attributes. Hellwig's algorithm, on the other hand, has shown that the redundant attributes do not correlate directly with such a standard statistical measure like Pearson correlation coefficient. Lack of the correlation might indicate non-linear dependency of some attributes with the decision class. Next, imbalance between the classes (212 and 55 elements, respectively) also does not help. To classify the instances from SPECT Heart data base with hight accuracy, quite large number of rules derived by application of rough sets was necessary [19]. It was also stressed [19] that the smaller class representing ill patients, consolidates 25 different clinical diseases. The difficulty with performing such an analysis is the lack of class examples – many categories have just a small number of examples, which renders most classifiers ineffective.

4 Conclusions

We have proposed a new method for feature selection for data bases for which the IFS model is justified. We have used the three term model of IFSs which made it possible to formulate a well motivated, very natural and understandable function being a foundation of the selection process. The method is transparent, simple from the point of view of calculation, fast, intuitively appealing and gives promising results. More, the attributes are ranked which is not always possible while using other methods of selecting the attributes. We plan to verify the proposed method using more data bases.

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. (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] Baldwin J.F., J. Lawry, T.P. Martin (1998), The Application of generalized Fuzzy Rules to Machine Learning and Automated Knowledge Discovery. Internationa Journal of Uncertainty, Fuzzyness and Knowledge-Based Systems, Vol. 6, No. 5, 459–487.

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

[7] 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., urada J. (Eds.): Artificial Intelligence and Soft Computing, Part I. Springer, Switzerland, 181-192.

[8] 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.

[9] Cios K.J., Kurgan Ł.A. UCI Machine Learning Repository databases. https://archive.ics.uci.edu/ml/datasets/spect+heart.

[10] Cios K.J., Kurgan Ł.. A. CLIP4: Hybrid inductive machine learning algorithm that generates inequality rules. Information Sciences 163 (2004) 37–83.

[11] 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.

[12] Hellwig Z. (1968) On the optimal choice of predictors. Study VI in Z. Gostkowski (ed.): Toward a system of quantitative indicators of components of human resources development; Paris: UNESCO, 1968.

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

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

[15] Jolliffe I.T. (2002) Principal Component Analysis. (2ed.) Springer.

[16] Kurgan L. A., K.J. Cios K.J., Tadeusiewicz R., Ogiela M., Goodenday L.S. Knowledge discovery approach to automated cardiac SPECT diagnosis. Artificial Intelligence in Medicine 23 (2) (2001) 149–169.

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

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

[19] S. el Rafaie, Abdel-Badeeh M. Salem, Revett K. (2012) On the use of SPECT imaging datasets for automated classification of ventricular heart disease. The 8th International Conference on INFOrmatics and Systems (INFOS2012) - 14–16 May. https://ieeexplore.ieee.org/document/ /6236549/?part=1

[20] 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.

[21] Quinlan J.R. (1986) Induction of decision trees. Machine Learning, 1, 81–106.

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

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

[24] 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.

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

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

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

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

[29] 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.

[30] 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.

[31] 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.

[32] 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.

[33] Szmidt E., Kacprzyk J. (2018) A New Approach to Hellwigs Method of Data Reduction for Atanassovs Intuitionistic Fuzzy Sets In: J. Medina et al. (Eds.): IPMU 2018, CCIS 855, pp. 553–564, 2018.

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

[35] 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.

[36] 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.

[37] Yager R.R. (1979) Level sets for membership evaluation of fuzzy subsets. Tech. Rep. RRY-79-14, Iona Colledge, New York. Also in: R.Yager, Ed., Fuzzy Set and Possibility Theory – Recent Developments. Pergamon Press, Oxford 1982, 90–97.

[38] Yamada K. (2001) Probability–Possibility Transformation Based on Evidence Theory. Proc. IFSA–NAFIPS'2001, 70–75.

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