Selection of the attributes
in intuitionistic fuzzy models

Eulalia Szmidt[1,2] and Janusz Kacprzyk[1,2]

[1] Systems Research Institute, Polish Academy of Sciences
ul. Newelska 6, 01-447 Warsaw, Poland, and
[2] Warsaw School of Information Technology
ul. Newelska 6, 01-447 Warsaw, Poland
e-mails: {szmidt,kacprzyk}@ibspan.waw.pl

To Professor Beloslav Riečan (1936–2018)

Abstract: We present a novel method of attribute selection for data bases which are expressed via intuitionistic fuzzy sets (IFSs, for short). We use the three term representation of the IFSs which makes it possible to construct a transparent and justified function that makes it possible to select attributes for widely understood decision making, e.g., for classification tasks. We test the proposed method using a well known example from literature. The results obtained are compared with other methods.
Keywords: Intuitionistic fuzzy sets, Three term representation of IFSs, Selection of attributes.
2010 Mathematics Subject Classification: 03E72, 34Gxx.

1   Introduction

The problem of model dimensionality reduction has been investigated for long and in spite of new approaches constantly proposed (with their pros and cons), the process continues as there is no overall "best method". We have two possible approaches to the model dimensionality reduction. First, 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 intuitionistic fuzzy sets (IFSs for short).

The intuitionistic fuzzy sets [1–3], are a very convenient tool for the modeling of systems in the presence of a lack of knowledge which is crucial for decision making and the same time difficult to foresee. The IFSs, being an extension of Zadeh's fuzzy sets [28], can make it possible to take into account a lack of knowledge by making use of the so-called hesitation margin or intuitionistic fuzzy index.

However, the IFS models can again 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. The well known Principal Component Analysis (PCA) for the IFSs [14, 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 make 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 test the proposed method using an example well known from the literature. The results are compared with those obtained by other methods of dimensionality reduction.

2 A brief introduction to the IFSs

One of the possible generalizations of a fuzzy set in X [28] 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 [1–3] 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 [15] 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 [2]

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

a *hesitation margin* of $x \in A$ which expresses a lack of knowledge of whether x belongs to A or not (cf. [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 [16, 17, 19], entropy (Szmidt and Kacprzyk [18, 20]), similarity (Szmidt and Kacprzyk [21]) for the IFSs, etc. i.e., the measures that play a crucial role in virtually all information processing tasks (Szmidt [14]).

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 [25–27]), the classification applying intuitionistic fuzzy trees (cf. Bujnowski [6]), group decision making (e.g., [4]), genetic algorithms [12], negotiations, voting and other situations (cf. Szmidt and Kacprzyk papers).

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

3 Three term representation of the IFSs as a foundation for selecting attributes

In [15] 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 [15] 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 (5), non-memberships (6), and hesitancy 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.

The most relevant attributes should be most discriminative. For a specific intuitionistic fuzzy attribute A_k it means that its average intuitionstic fuzzy index (7) 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:

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

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 (8) are:

1. $0 \leq f(A_k) \leq 1$.

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 (8), and its contour plot.
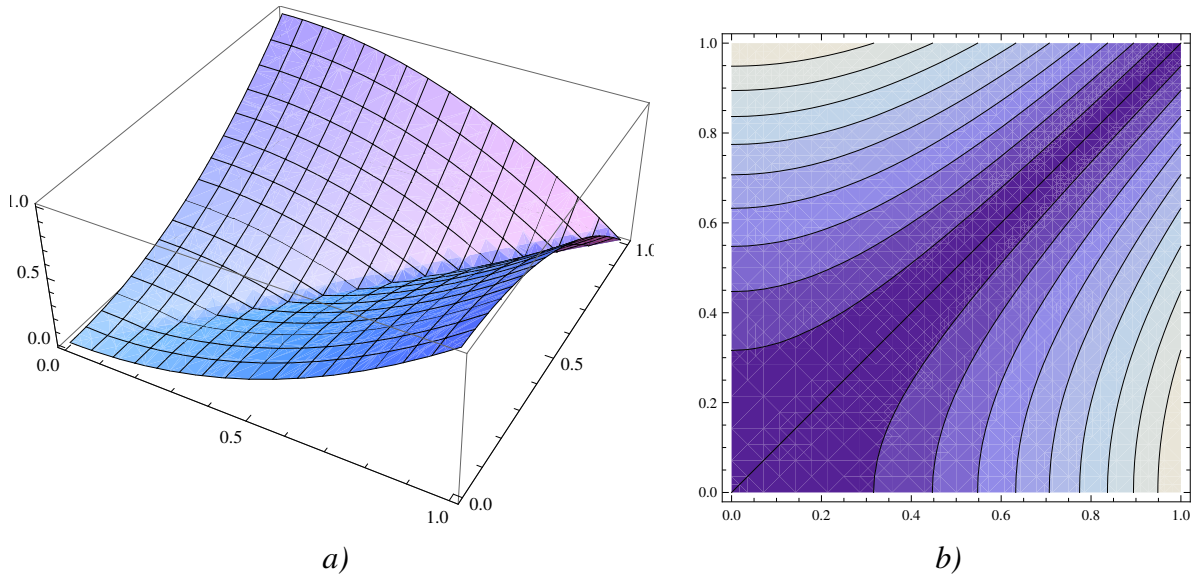

a)   b)

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

It is worth noticing that the components in (8) are independent, i.e. $|\overline{\mu}_k - \overline{\nu}_k|$ is independent on π since the shape of $|\overline{\mu}_k - \overline{\nu}_k|$ is always the same in spite of π.

From (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 , $k = 1, \dots, K$.

Repeating (9) $K-1$ times we can order all K attributes: from the most to the least discriminative.

4 Results

As an illustration how the proposed method works, we will recall a well known problem formulated by Quinlan [13] but expressed in terms of the IFSs. The 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).

In Quinlan's example [13] (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 [13]:

- **outlook**, with values {sunny, overcast, rain},

- **temperature**, with values {cold, mild, hot},

- **humidity**, with values {high, normal}, and

- **windy**, with values {true, false},

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 [13]

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 [22]) and we only present here the final results.

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

Next, making use of the IFS model (Table 2) we compute an average characteristic of the attributes. The results are in Table 3.

1	2	3	4	5	6
Attribute	$\overline{\pi}$	$\overline{\mu}$	$\overline{\nu}$	$|\overline{\mu} - \overline{\nu}|$	$(1 - \overline{\pi})(|\overline{\mu} - \overline{\nu}|$
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. Characteristic of the "Saturday Morning" attributes

In the last column of Table 3 there are values obtained from (8) for each attribute. Making use of (9) we can order the attributes (the bigger value from (8) the better), namely: Outlook, Humidity, Windy, Temperature. We can notice that that the values (8) are not too big what is the result of both substantial values of hesitation margins (second column of Table 3) and not very big differences between $\overline{\mu}$ and $\overline{\nu}$.

It is worth noticing that Quinlan [13] obtained 100% classification accuracy, and the optimal solution (the minimal possible ID3 tree) using the first three attributes as pointed out by our method (Outlook, Humidity, Windy).

Another selection of the attributes obtained via Hellwig's method [8] adapted to IFS data (cf. Szmidt and Kacprzyk [24]) and verified using the Quinlan's example [24] 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 [9–11] in the sense of mean-square error.

After performing the PCA adapted to the data expressed via the IFSs (cf. Szmidt and Kacprzyk [23]), 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 (9) presented here (PCA as a method of reduction points out a combination of attributes not the initial attributes).

Clearly, our example is just for illustration as feature reduction makes sense for large problems (very many features) and then the selection of the attributes is usually considerable and very welcome.

5 Conclusions

We have proposed a novel 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 very natural and understandable function being a foundation of the selection process. The method is transparent, simple from the point of view of calculation, intuitively appealing and gives promising results. We plan to verify the proposed method on bigger data bases.

References

[1] Atanassov, K. (1983). Intuitionistic fuzzy sets. *VII ITKR Session*, Sofia 20–23 June, 1983 (Central Sci. and Techn. Library, Bulg. Academy of Sciences, 1984). (in Bulgarian). Reprinted: *Int. J. Bioautomation*, 2016, 20(S1), S1–S6 (in English).

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

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

[4] Atanassova, V. (2005). Strategies for decision making in the conditions of intuitionistic fuzziness. *Reusch B. (eds) Computational Intelligence, Theory and Applications*. Advances in Soft Computing, vol 33. Springer, Berlin, Heidelberg, 263–269.

[5] Baldwin, J. F., Lawry, J., & Martin, T. P. (1998). The Application of generalized Fuzzy Rules to Machine Learning and Automated Knowledge Discovery. *Internationa Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 6(5), 459–487.

[6] Bujnowski, P., Szmidt, E., & Kacprzyk, J. (2014). Intuitionistic Fuzzy Decision Trees - a new Approach. *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. *Proc. of IPMU'2006*, 2383–2390.

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

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

[12] Roeva, O., & Michalikova, A. (2013). Generalized net model of intuitionistic fuzzy logic control of genetic algorithm parameters. *Notes on Intuitionistic Fuzzy Sets*, 19(2), 71–76.

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

[14] Szmidt, E. (2014). *Distances and Similarities in Intuitionistic Fuzzy Sets*. Springer.

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

[16] Szmidt, E., & Kacprzyk, J. (1997). On measuring distances between intuitionistic fuzzy sets, *Notes on Intuitionistic Fuzzy Sets*, 3(4), 1–13.

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

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

[19] Szmidt, E. & Kacprzyk, J. (2006). Distances Between Intuitionistic Fuzzy Sets: Straightforward Approaches may not work. *Proc. of IEEE Intelligent Systems'06*, 716–721.

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

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

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

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

[24] Szmidt, E. & Kacprzyk, J. (2018). A New Approach to Hellwig's Method of Data Reduction for Atanassov's Intuitionistic Fuzzy Sets. *Medina, J. et al. (Eds.): IPMU 2018, CCIS 855*, 553–564.

[25] Szmidt, E., & Kukier, M. (2006). Classification of Imbalanced and Overlapping Classes using Intuitionistic Fuzzy Sets. *Proc. of IEEE Intelligent Systems'06*, London, 722–727.

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

[27] Szmidt, E., & Kukier, M. (2008a). Atanassov's intuitionistic fuzzy sets in classification of imbalanced and overlapping classes. *Chountas, P., Petrounias, I., Kacprzyk, J. (Eds.): Intelligent Techniques and Tools for Novel System Architectures*. Springer, Berlin Heidelberg 2008, 455–471.

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