Intuitionistic Fuzzy Decision Trees - A New Approach

Paweł Bujnowski, Eulalia Szmidt, and Janusz Kacprzyk

Systems Research Institute, Polish Academy of Sciences,
ul. Newelska 6, 01–447 Warsaw, Poland
Warsaw School of Information Technology, ul. Newelska 6, 01-447 Warsaw, Poland
pbujno@gmail.com, {szmidt,kacprzyk}@ibspan.waw.pl

Abstract. We present here a new classifier called an intuitionistic fuzzy decision tree. The performance of the new algorithm is illustrated by providing an analysis of well known benchmark data. The results are compared to some other well known classification algorithms.

1 Introduction

Decision trees, with their well known advantages, are very popular classifiers which recursively partition a space of instances (observations). Following the source Quinlan the ID3 algorithm [21], many other approaches have been developed along that line (cf. [25]).

Classical (crisp) decision trees were extended to fuzzy decision trees which turned out to be more stable, and effective method to extract knowledge in uncertain classification problems (Janikow [16], Olaru et al. [20], Yuan and Shaw [38], Marsala [18], [19]).

The next natural step is to take advantages of the intuitionistic fuzzy sets introduced by Atanassov [1], [2], [3] (A-IFSs for short) while building the trees.

In this paper we propose a new intuitionistic fuzzy decision tree classifier. The data is expressed by means of intuitionistic fuzzy sets. Also the measures constructed for the intuitionistic fuzzy sets are applied while making decisions how to split a node while expanding the tree. The intuitionistic fuzzy tree proposed here is an extension of the fuzzy ID3 algorithm [6].

The potential of the new algorithm is illustrated by providing an analysis of well known benchmark data. The results are compared to other commonly used algorithms.

2 A Brief Introduction to A-IFSs

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

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

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

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

L. Rutkowski et al. (Eds.): ICAISC 2014, Part I, LNAI 8467, pp. 181–192, 2014.

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

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

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. (An approach to the assigning memberships and non-memberships for A-IFSs from data is proposed by Szmidt and Baldwin [26]).

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

An additional concept for each A-IFS in X, that is not only an obvious result of (2) and (3) but which is also relevant for applications, we will call (Atanasov [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. 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 [27], [28], [30], entropy (Szmidt and Kacprzyk [29], [31]), similarity (Szmidt and Kacprzyk [32]) for the A-IFSs, etc. i.e., the measures that play a crucial role in virtually all information processing tasks.

Hesitation margins turn out to be relevant for applications - in image processing (cf. Bustince et al. [14], [13]) and classification of imbalanced and overlapping classes (cf. Szmidt and Kukier [33], [34], [35]), group decision making, negotiations, voting and other situations (cf. Szmidt and Kacprzyk papers).

3 Intuitionistic Fuzzy Decision Tree - New Algorithm Description

The intuitionistic fuzzy decision tree proposed here has its roots in the soft decision tree introduced by Baldwin et al. [6] which follow the source *ID3* tree introduced by Quinlan [21].

We consider numeric attributes but the methods presented here can be also easily applied to the nominal attributes (the algorithm is even simpler then). We use here intuitionistic fuzzy sets for data representation. More, the new idea of deriving intuitionistic fuzzy sets in each node was applied as potentially giving the most accurate results.

Splitting the nodes is the most important step in generating a decision tree. The step demands to point out the best attributes for splitting. Proper picking up the attributes influences accuracy of a decision tree, and its interpretation properties. In the tree presented here intuitionistic fuzzy entropy was used (Szmidt and Kacprzyk[29]) as a counterpart of “information gain” [21].

In the next sections the most important components of the algorithm are described.

3.1 Fuzzy Partitions of the Attribute Values (granulation)

The idea of a universe partition (granulation), i.e., replacing a continuous domain with a discrete one has been extended to fuzzy sets by Ruspini [23]. The idea was used here

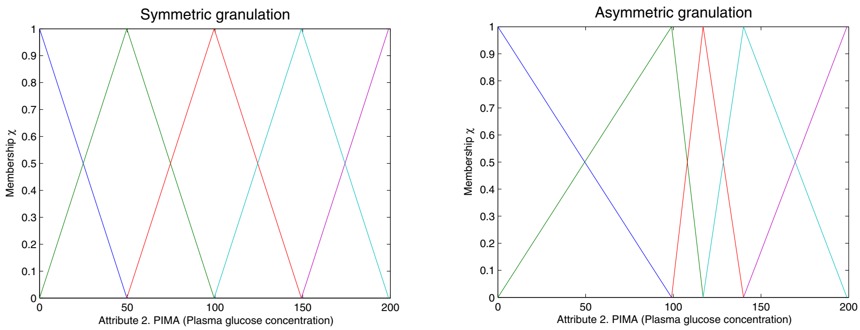

Fig. 1. Illustration of symmetric fuzzy partitioning, and asymmetric fuzzy partitioning (on attribute 2 “Plasma glucose concentration” of benchmark “Pima Diabetes” with 5 fuzzy sets)

to partition a universe of each attribute by introducing a set of triangular fuzzy sets such that for any attribute value the sum of memberships of the partitioning fuzzy sets is 1.

More formally, the membership $\chi_{j,k}(o_{ij})$ of the i-th observation (instance) o_{ij} in respect to the j-th attribute to the triangular fuzzy sets k and $k+1$ (where $k = 1, \ldots, p$) is:

$$\chi_{j,k}(o_{ij}) + \chi_{j,k+1}(o_{ij}) = 1, \quad k = 1, \ldots, p-1, \tag{5}$$

and for the j-th attribute A_j we have $o_{ij} \in A_j,\ i = 1, \ldots, n,\ j = 1, \ldots, m.$
In other words, the sum of the membership values for an observation o_{ij} is one (the sum results from only two neighboring fuzzy sets).

Remark. Here, for the purpose of granulation we use symbol χ for the membership values so to make a difference between membership values resulting from the attribute granulation (χ) and the membership values of the intuitionistic fuzzy sets μ.

We use here symmetric, evenly spaced triangular fuzzy sets (symmetric fuzzy partitions), and asymmetric, unevenly spaced triangular fuzzy sets (asymmetric fuzzy partitions such that each partition contains equal number of data points) [4,23]. In Fig. 1 an example is shown of symmetric fuzzy partitioning (symmetric granulation), and asymmetric fuzzy partitioning (asymmetric granulation). The two kinds of partitioning are illustrated on attribute 2 of the “PIMA Diabetes” problem with 5 fuzzy sets. Fuzzy partitioning (triangular fuzzy sets) is a starting point to assign nodes in a soft ID3 decision tree - cf. Fig. 2.

3.2 Fuzzy *ID3* Algorithm

In this section we present a fuzzy generalization of *ID3* algorithm [6].

Consider the following database

$$T = \{o_i =< o_{i,1}, \ldots, o_{i,m} > \mid i = 1, \ldots, n\}, \tag{6}$$

where $o_{i,j}$ is a value of the j-th attribute $A_j, j = 1, \ldots, m$, for the i-th instance. We assume that $o_{i,j}$ are crisp.

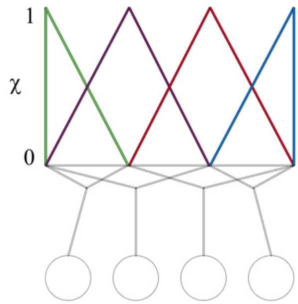

Fig. 2. Fuzzy partitioning as a starting point to constructing nodes in a soft ID3 tree

At the beginning of generating a fuzzy *ID3* decision tree from data, its root contains all the instances (top down approach). Each node is split by partitioning its instances. A node becomes a leaf if all the attributes are used in the path considered or if all its instances are from a unique class.

Splitting the nodes in a decision tree can be represented by the rules. Assume that P_j is a partition set of the attribute space Ω_j ($j = 1, \ldots, m$), and that partition of each attribute is via triangular fuzzy sets. Let $P_{\chi_{j,k}} \in P_j$ be the k-th partitioning fuzzy set expressed by a triangle membership function $\chi_{j,k}$ being a component of the partition of the j-th attribute. The following rule expresses conjunction of the fuzzy conditions along the path from the root to a tree node

$$B \equiv P_{\chi_{j_1}} \wedge \cdots \wedge P_{\chi_{j_N}} \tag{7}$$

where $P_{\chi_{j_r}}$ are triangular fuzzy sets, and its set of indexes represented by the subsequence (j_r) is in a considered rule a result of pointing up a pair: (1) a unique attribute numbers j, and (2) one from the k triangle fuzzy sets for each attribute partitioning. Formula (7) expresses a conjunction of the conditions which are to be fulfilled for an instance o_i so that it were present in a considered node. Database $T = \{o_i,\ i = 1, \ldots, n\}$ generates a *support* for B (7) given as:

$$w(B) = \sum_{i=1}^{n} \prod_{j_r} Prob(P_{\chi_{j_r}} | o_i) \tag{8}$$

where $Prob(P_{\chi_{j_r}} | o_i)$ is a probability defined on the fuzzy set $P_{\chi_{j_r}}$ provided the observation o_i. It is easily calculated using the membership function $\chi_{j_r(o_i)}$.

Let $\{C_l,\ l = 1, \ldots, h\}$ denotes a set of decision classes. Formula (8) is also used for generating support for a given decision class, e.g., C_x in a given node, namely

$$Prob(C_x | B) = \frac{w(C_x \wedge B)}{\sum_{l=1}^{h} w(C_l \wedge B)} = \frac{w(C_x \wedge B)}{w(B)}. \tag{9}$$

Splitting a node (starting from a root) is related to the attributes' abilities evaluation to generate a next level with the child nodes. A potential possibility of an attribute A for producing child nodes $A_s, s = 1, \ldots p$ is tested by calculating its classical entropy:

$$I(A_s) = -\sum_{l=1}^{h} Prob(C_l|A_s)\ log(Prob(C_l|A_s)),\ s = 1, \dots p., \tag{10}$$

The common entropy for an attribute A is the following weighted mean value:

$$I(A) = \frac{\sum_{s=1}^{p} w(A_s) \cdot I(A_s)}{\sum_{s=1}^{p} w(A_s)} \tag{11}$$

In (10) and (11) it has been assumed that A_s represents a rule from the root to the s-th child node.

Using the presented above formulas makes it possible to generate the nodes in a fuzzy *ID3* tree [6].

3.3 Deriving Intuitionistic Fuzzy Sets from Data

We will present now a modification of the soft *ID3* approach (Section 3.2) by using intuitionistic fuzzy sets.

Let assume that an attribute A, splitting a node into the child nodes $A_s, s = 1, \dots p$, is tested. For simplicity we assume that only two decision classes C^+ and C^- are considered. Support for these classes in each node is

$$\begin{array}{l} for\ class\ C^+ : w(C^+ \wedge A_1), w(C^+ \wedge A_2), \cdots, w(C^+ \wedge A_p) \\ for\ class\ C^- : w(C^- \wedge A_1), w(C^- \wedge A_2), \cdots, w(C^- \wedge A_p). \end{array} \tag{12}$$

Independently for each class their frequencies for the verified splitting are calculated (proportions between support of a class in the child nodes and its cardinality in the parent node)

$$\begin{array}{l} p(C^+|A_s) : \frac{w(C^+ \wedge A_1)}{w(C^+ \wedge A)}, \frac{w(C^+ \wedge A_2)}{w(C^+ \wedge A)}, \cdots, \frac{w(C^+ \wedge A_p)}{w(C^+ \wedge A)} \\ p(C^-|A_s) : \frac{w(C^- \wedge A_1)}{w(C^- \wedge A)}, \frac{w(C^- \wedge A_2)}{w(C^- \wedge A)}, \cdots, \frac{w(C^- \wedge A_p)}{w(C^- \wedge A)}. \end{array} \tag{13}$$

Having the relative frequencies $p(C^+|A_i)$ and $p(C^-|A_i)$ (13), we use the algorithm given in [5,6] to construct independently fuzzy sets representing the classes C^+, and C^-. The fuzzy sets obtained for C^+, and C^- are abbreviated Pos^+ and Pos^-, respectively. In the fuzzy *ID3* tree [6] the fuzzy sets $Pos^+(A_s)$ and $Pos^-(A_s),\ s = 1, \dots, p$ are tested by a classical entropy (10) - (11) to assess the attributes.

In the algorithm proposed here we use the fuzzy model (expressed by Pos^+ and Pos^-) to construct intuitionistic fuzzy model (details are presented in Szmidt and Baldwin [26]). Intuitionistic fuzzy model of the data in the child nodes $A_s,\ s = 1, \dots p$ (due to the algorithm in [26]) is expressed by the following intuitionistic fuzzy terms

$$\begin{array}{l} \pi(A_s) = Pos^+(A_s) + Pos^-(A_s) - 1 \\ \mu(A_s) = Pos^+(A_s) - \pi(A_s) \\ \nu(A_s) = Pos^-(A_s) - \pi(A_s). \end{array} \tag{14}$$

This way each child node s is described by the following intuitionistic fuzzy set

$$< A_s, \mu(A_s), \nu(A_s), \pi(A_s) >,\ s = 1, \dots, p \tag{15}$$

where μ describes support for the class C^+; ν describes support for the class C^-; π expresses lack of knowledge concerning μ and ν.

An instance o_i characteristic at node A_s can be expressed as well in terms of intuitionistic fuzzy sets

$$\chi_{A_s}(o_i) \cdot < \mu(A_s), \nu(A_s), \pi(A_s) >, \; i = 1, \ldots, n,$$

where χ_{A_s} is a membership function at node A_s expressed by the product in (8). Having in mind the property (5) we can obtain full information value of an instance o_i while partitioning A and obtaining in result the child nodes $\{A_s, \; s = 1, \ldots, p\}$:

$$\chi_{A_s}(o_i) \cdot < \mu(A_s), \nu(A_s), \pi(A_s) > + \chi_{A_{s+1}}(o_i) \cdot < \mu(A_{s+1}), \nu(A_{s+1}), \pi(A_{s+1}) > . \tag{16}$$

Both (15) and (16) may be used (alternatively) in the algorithm proposed for assessing and choosing the attributes while splitting the nodes in the intuitionistic fuzzy decision tree.

3.4 Selection of an Attribute to Split a Node

In the process of expanding a tree – a crisp, fuzzy or intuitionistic fuzzy tree, the crucial step is splitting a node into children nodes. To split a node an attribute is selected on the basis of its "information gain". Different measures may be used to assess "information gain". We use here an intuitionistic fuzzy measure – intuitionistic fuzzy entropy [29].

Intuitionistic fuzzy entropy $E(x)$ of an intuitionistic fuzzy element $x \in A$ is given as [29]:

$$E(x) = \frac{min\{l_{IFS}(x, M), l_{IFS}(x, N)\}}{max\{l_{IFS}(x, M), l_{IFS}(x, N)\}}, \tag{17}$$

where M, N are the intuitionistic fuzzy elements ($< \mu, \nu, \pi >$) fully belonging (M) or fully not belonging (N) to a set considered

$$M =< 1, 0, 0 >$$
$$N =< 0, 1, 0 >,$$

$l_{IFS}(\cdot, \cdot)$ is the normalized Hamming distance [28,30]:

$$l_{IFS}(x, M) = \tfrac{1}{2}(|\mu_x - 1| + |\nu_x - 0| + |\pi_x - 0|)$$

$$l_{IFS}(x, N) = \tfrac{1}{2}(|\mu_x - 0| + |\nu_x - 1| + |\pi_x - 0|).$$

It is also possible to use other intuitionistic fuzzy measures to evaluate the attributes (cf. [36], [37]) but due to the space limitation here we discuss entropy only.

Intuitionistic fuzzy entropy of an intuitionistic fuzzy set with n elements: $X = \{x_1, \ldots, x_n\}$ is [29]:

$$E(X) = \frac{1}{n} \sum_{i=1}^{n} E(x_i). \tag{18}$$

To compute intuitionistic fuzzy entropy $E(A_s)$ (17) in a child node $A_s,\ s = 1, \ldots, p$, we make use of the intuitionistic fuzzy representations (12)–(15) of the possible child nodes derived while testing attribute A.

Total intuitionistic fuzzy entropy of an attribute A is abbreviated $E(W_A)$ whereas entropy of a child node – $E(A_s)$. Total intuitionistic fuzzy entropy of A is a sum of the weighted intuitionistic fuzzy entropy measures of all the child nodes $A_s,\ s = 1, \ldots, p$, with the weights reflecting supports (cardinalities) of the nodes:

$$E(W_A) = \frac{\sum_{s=1}^{p} w(A_s) E(A_s)}{\sum_{s=1}^{p} w(A_s)}. \tag{19}$$

Instead of using (19) we may calculate $E(W_A)$ by applying a weighted intuitionistic fuzzy representation of each instance o_i (16) while partitioning an attribute A. Next, using (18), a total intuitionistic fuzzy entropy is calculated for a chosen attribute. This method was applied in the numerical experiments (cf. Section 4).

An attribute for which total intuitionistic fuzzy entropy is minimal is selected for splitting a node.

A flowchart representing a process of generating intuitionistic fuzzy decision tree is in Fig. 3.

3.5 Classification of the Instances

A leaf in a soft tree is described by a proportion of the classes considered. A single instance usually belongs to several leaves. In result we need aggregated information about total degree of membership of a single observation to each class.

To classify the instances we use here measure *SUM* which is a sum of the products of the instance membership values at leafs and support for a class considered in these leafs [6]. Total support of the observation $o_i,\ i = 1, \ldots, n$, for a class C is:

$$supp(C|o_i)_{SUM} = \sum_{j=1}^{L} supp(C|T_j) \cdot \chi(T_j|o_i), \tag{20}$$

where $\{T_j : j = 1, \ldots, L\}$ is a set of the leafs; L is the number of the leafs; $supp(C|T_j)$ is a support of the classes considered in the j-th leaf; $\chi(T_j|o_i)$ is a membership value of the observation o_i (it is a result of the partitioning of the universe attributes), different for each leaf, fulfilling: $\sum_{j=1}^{L} \chi(T_j|o_i) = 1$.

4 Results of the Numerical Experiments

We have verified classification abilities of the new intuitionistic fuzzy decision tree with other well known algorithms.

The following measures were used to compare the behavior of the classifiers compared:

– total proper identification of the instances belonging to the classes considered,
– the area under ROC curve [15].

Behavior of the intuitionistic fuzzy decision tree proposed here was compared especially with other decision trees, namely:

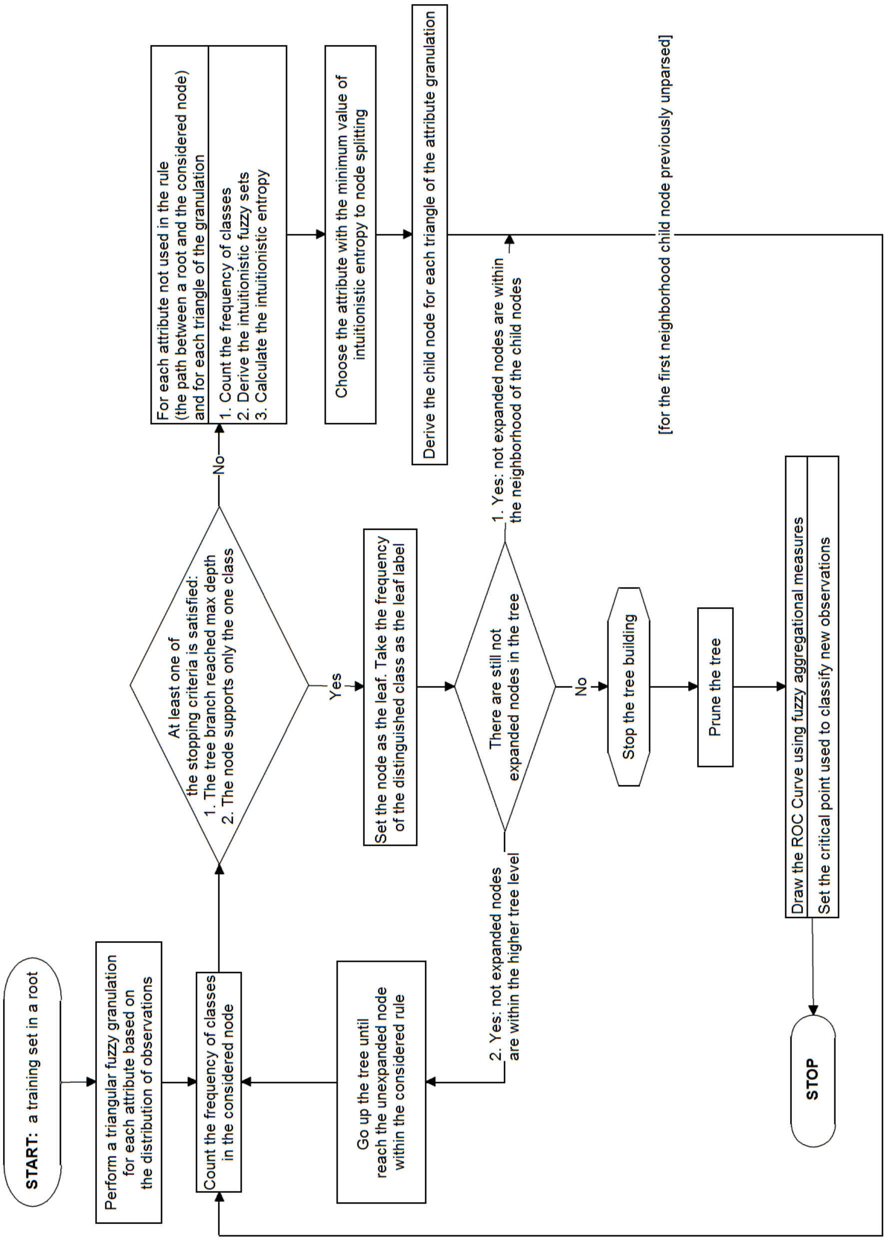

Fig. 3. A flowchart representing a process of generating intuitionistic fuzzy tree

Table 1. “Sonar” benchmark data – comparison of the intuitionistic fuzzy decision tree and other classifiers

	Classification accuracy ($\bar{x} \pm \sigma$) [%]	
Algorithm	**for both classes**	**AUC ROC**
intuitionistic fuzzy tree (asym)	80.80 ± 7.76 $(*)$	89.81 ± 5.66 $(*)$
$RandomForest$	80.41 ± 8.80	89.53 ± 7.58
$MultilayerPerceptron$	81.61 ± 8.66	88.48 ± 7.31
pruned intuitionistic fuzzy tree (asym)	78.63 ± 7.89	86.92 ± 6.29 $(-)$
LMT	76.27 ± 9.62 $(-)$	84.15 ± 8.55 $(-)$
$NBTree$	77.07 ± 9.65 $(-)$	83.10 ± 9.89 $(-)$
SDT $(refitting)$	$73.28 \pm b.d.$ $(b.d.)$	$b.d.$
SDT $(backfitting)$	$72.56 \pm b.d.$ $(b.d.)$	$b.d.$
$Logistic$	72.47 ± 8.90 $(--)$	80.02 ± 8.78 $(--)$
$J48$ $(unpruned\, C4.5)$	73.42 ± 9.36 $(-)$	79.37 ± 10.83 $(--)$
$J48$ $(pruned\, C4.5)$	73.61 ± 9.34 $(-)$	79.31 ± 10.80 $(--)$

- ***J48*** – implementation of the crisp tree proposed by Quinlan *C4.5* ([22]),
- ***LMT*** (*Logistic Model Tree*) – a hybrid tree building the logistic models at the leaves ([17]),
- ***NBTree*** – hybrid decision tree building the Bayes classifiers at the leaves,
- ***RandomForest*** – here consisting of 10 decision trees which nodes are generated on the basis of a random set of attributes ([10]).

Besides the trees, also neural networks (***MultilayerPerceptron***), and logistic regression (***Logistic***) were used for the evaluation. The evaluation of the above algorithms was performed using WEKA (http://www.cs.waikato.ac.nz/ml/weka/). Next, the results obtained by Olaru and Wehenkel [20] using *Soft Decision Trees (SDT)* are compared.

We present here the results obtained by intuitionistic fuzzy decision tree for “Sonar” benchmark data (http://archive.ics.uci.edu/ml/datasets.html). The dataset contains 208 instances, 60 numerical attributes, 2 classes (111 – metal cylinder, and 97 instances – rocks). We use simple cross validation method with 10 experiments of 10-fold cross validation (giving 100 trees). For each experiment an average value of the accuracy measures, and of their standard deviations were calculated. So to compare an average accuracy of the new intuitionistic fuzzy decision tree with other classifiers, *t-Student* test was used (Table 1). One minus in Table 1 means that the (worse) result was obtained by a classifier while using classical *t-Student* test, two minuses mean using corrected *t-Student* test (for cross validation).

Results obtained (Table 1 – accuracy, and Fig. 4 – ROC curves) show that the intuitionistic fuzzy decision tree is the best concerning the area under ROC curve, and the second one in respect of accuracy (a little worse than *Multilayer Perceptron*). In other words, the new intuitionistic fuzzy decision tree turned out a better classifier for “Sonar” benchmark data than other crisp and soft decision trees, even slightly better than *Random Forest*, and almost as effective as *Multilayer Perceptron*.

Surprisingly enough, just for the “Sonar” benchmark data, the proposed classifier turned out to be worse than the simplest k-nearest neighbor classifier. Due to space

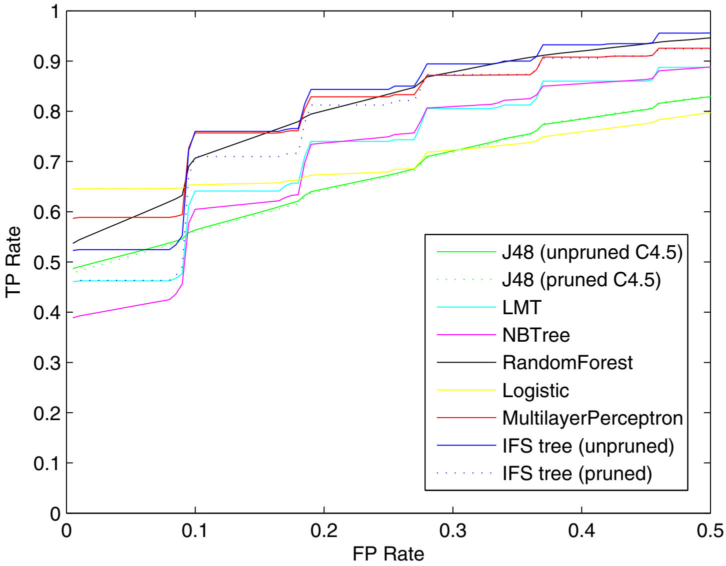

Fig. 4. "Sonar" benchmark data – comparison of the areas under ROC curves for the intuitionistic fuzzy decision tree and other classifiers (TP – True Positive, FP – False Positive rates [15])

Table 2. Comparison of the accuracy of k-NN classifier, trees.J48 and the intuitionistic fuzzy decision tree for chosen benchmark data

	Classification accuracy ($\bar{x} \pm \sigma$) [%]			
Data set	**k-NN for k=1**	**k-NN for k=3**	**trees.J48 (pruned)**	**IFS tree**
$PIMA$	70.62 ± 4.67	73.86 ± 4.55	74.49 ± 5.27	$\mathbf{75.72 \pm 4.37}$
$Sonar$	$\mathbf{86.17 \pm 8.45}$	$\mathbf{83.76 \pm 8.51}$	73.61 ± 9.34	80.80 ± 7.76
$Ionosphere$	87.10 ± 5.12	86.02 ± 4.31	89.74 ± 4.38	$\mathbf{90.36 \pm 4.50}$
$Wine$	95.12 ± 4.34	95.85 ± 4.19	93.20 ± 5.90	$\mathbf{97.88 \pm 3.53}$
$Glass$	70.30 ± 8.96	69.84 ± 8.61	67.61 ± 9.26	$\mathbf{75.16 \pm 6.21}$
$Iris$	95.40 ± 4.80	95.20 ± 5.11	94.73 ± 5.30	$\mathbf{96.20 \pm 4.37}$

limitation we do not present a detailed comparison of the proposed classifier for other data sets (as for the "Sonar" data – Table 1), but results of the experiments with the k-nearest neighbor classifier (Table 2) are added for several other benchmark data sets, namely: "PIMA", "Ionosphere", "Wine", "Glass", "Iris" (http://archive.ics.uci.edu/ml/datasets.html). It is easy to notice that the proposed classifier produces more accurate results than the k-nearest neighbor classifier (the "Sonar" benchmark data is an exception). In other words, the proposed classifier may be not the best solution for all possible data sets, but no other classifier can be, for obvious reasons! However, for the data sets presented in Table 2 it turned out to be usually better, and certainly not worse than the classifiers presented in Table 1. In addition, as a tree type classifier, it can be a properer, if not the best choice, in many applications when comprehensibility and transparency to the human being is relevant.

5 Conclusions

We have proposed a new intuitionistic fuzzy decision tree which is an extension of the fuzzy ID3 decision tree algorithm. The tree proposed was tested on a well known benchmark examples. The results are very encouraging.

References

1. Atanassov, K.: Intuitionistic Fuzzy Sets. VII ITKR Session. Sofia (June 1983) (Deposed in Central Sci.-Techn. Library of Bulgarian Academy of Sciences, 1697/84)
2. Atanassov, K.: Intuitionistic Fuzzy Sets: Theory and Applications. Springer (1999)
3. Atanassov, K.: On Intuitionistic Fuzzy Sets Theory. Springer (2012)
4. Baldwin, J.F., Karale, S.B.: Asymmetric Triangular Fuzzy Sets for Classification Models. In: Palade, V., Howlett, R.J., Jain, L.C. (eds.) KES 2003. LNCS (LNAI), vol. 2773, pp. 364–370. Springer, Heidelberg (2003)
5. Baldwin, J.F., Lawry, J., Martin, T.P.: A mass assignment theory of the probability of fuzzy events. Fuzzy Sets and Systems 83, 353–367 (1996)
6. Baldwin, J.F., Lawry, J., Martin, T.P.: Mass Assignment Fuzzy ID3 with Applications. In: Unicom Workshop on Fuzzy Logic Applications and Future Directions, London (1997)
7. Bartczuk, Ł., Rutkowska, D.: A New Version of the Fuzzy-ID3 Algorithm. In: Rutkowski, L., Tadeusiewicz, R., Zadeh, L.A., Żurada, J.M. (eds.) ICAISC 2006. LNCS (LNAI), vol. 4029, pp. 1060–1070. Springer, Heidelberg (2006)
8. Benbrahim, H., Bensaid, A.: A comparative study of pruned decision trees and fuzzy decision trees. In: NAFIPS 2000, pp. 227–231 (2000)
9. Bezdek, J.C.: Pattern Recognition with Fuzzy Objective Function Algorithms. Kluwer Academic Publishers, Norwell (1981)
10. Breiman, L.: Random Forests. Machine Learning 45(1), 5–32 (2001)
11. Breiman, L., Friedman, J.H., Olsen, R.A., Stone, C.J.: Classification and Regression Trees. Wadsworth, Belmont (1984)
12. Bujnowski, P.: Using intuitionistic fuzzy sets for constructing decision trees in classification tasks. PhD dissertation, IBS PAN, Warsaw (2013) (in Polish)
13. Bustince, H., Mohedano, V., Barrenechea, E., Pagola, M.: Image thresholding using intuitionistic fuzzy sets. In: Atanassov, K., Kacprzyk, J., Krawczak, M., Szmidt, E. (eds.) Issues in the Representation and Processing of Uncertain and Imprecise Information. Fuzzy Sets, Intuitionistic Fuzzy Sets, Generalized Nets, and Related Topics. EXIT, Warsaw (2005)
14. Bustince, H., Mohedano, V., Barrenechea, E., Pagola, M.: An algorithm for calculating the threshold of an image representing uncertainty through A-IFSs. In: IPMU 2006, pp. 2383–2390 (2006)
15. Hand, D.J., Till, R.J.: A simple generalization of the area under the ROC curve for multiple class classification problems. Machine Learning 45, 171–186 (2001)
16. Janikow, C.Z.: Fuzzy Decision Trees: Issues and Methods. IEEE Transactions on Systems, Man, and Cybernetics 28(1), 1–14 (1998)
17. Landwehr, N., Hall, M., Frank, E.: Logistic Model Trees. Machine Learning 95(1-2), 161–205 (2005)
18. Marsala, C.: Fuzzy decision trees to help flexible querying. Kybernetika 36(6), 689–705 (2000)
19. Marsala, C., Bouchon-Meunier, B.: An adaptable system to construct fuzzy decision tree. In: NAFIPS 1999, pp. 223–227 (1999)

20. Olaru, C., Wehenkel, L.: A complete fuzzy decision tree technique. Fuzzy Sets and Systems, 221–254 (2003)
21. Quinlan, J.R.: Induction of decision trees. Machine Learning 1, 81–106 (1986)
22. Quinlan, J.R.: C4.5: Programs for Machine Learning. Morgan Kaufman Publishers, Inc., San Mateo (1993)
23. Ruspini, E.H.: A New Approach to Clustering. Information and Control 15, 22–32 (1969)
24. Rutkowski, L.: Artificial intelligence methods and techniques, pp. 237–307. PWN, Warszawa (2009) (in Polish)
25. Safavian, S.R., Landgrebe, D.: A survey of decision tree classifier methodology. IEEE Trans. Systems Man Cybernet. 21, 660–674 (1991)
26. Szmidt, E., Baldwin, J.: Intuitionistic Fuzzy Set Functions, Mass Assignment Theory, Possibility Theory and Histograms. In: 2006 IEEE WCCI, pp. 237–243 (2006)
27. Szmidt, E., Kacprzyk, J.: On measuring distances between intuitionistic fuzzy sets. Notes on IFS 3(4), 1–13 (1997)
28. Szmidt, E., Kacprzyk, J.: Distances between intuitionistic fuzzy sets. Fuzzy Sets and Systems 114(3), 505–518 (2000)
29. Szmidt, E., Kacprzyk, J.: Entropy for intuitionistic fuzzy sets. Fuzzy Sets and Systems 118, 467–477 (2001)
30. Szmidt, E., Kacprzyk, J.: Distances Between Intuitionistic Fuzzy Sets: Straightforward Approaches may not work. In: 3rd International IEEE Conference Intelligent Systems, IEEE IS 2006, London, pp. 716–721 (2006)
31. Szmidt, E., Kacprzyk, J.: Some problems with entropy measures for the Atanassov intuitionistic fuzzy sets. In: Masulli, F., Mitra, S., Pasi, G. (eds.) WILF 2007. LNCS (LNAI), vol. 4578, pp. 291–297. Springer, Heidelberg (2007)
32. Szmidt, E., Kacprzyk, J.: A New Similarity Measure for Intuitionistic Fuzzy Sets: Straightforward Approaches may not work. In: 2007 IEEE Conf. on Fuzzy Systems, pp. 481–486 (2007a)
33. Szmidt, E., Kukier, M.: Classification of Imbalanced and Overlapping Classes using Intuitionistic Fuzzy Sets. In: IEEE IS 2006, London, pp. 722–727 (2006)
34. Szmidt, E., Kukier, M.: A New Approach to Classification of Imbalanced Classes via Atanassov's Intuitionistic Fuzzy Sets. In: Wang, H.-F. (ed.) Intelligent Data Analysis: Developing New Methodologies Through Pattern Discovery and Recovery, pp. 85–101. Idea Group (2008)
35. Szmidt, E., Kukier, M.: Atanassov's intuitionistic fuzzy sets in classification of imbalanced and overlapping classes. In: Chountas, P., Petrounias, I., Kacprzyk, J. (eds.) Intelligent Techniques and Tools for Novel System Architectures. SCI, vol. 109, pp. 455–471. Springer, Heidelberg (2008)
36. Szmidt, E., Kacprzyk, J., Bujnowski, P.: Measuring the Amount of Knowledge for Atanassov's Intuitionistic Fuzzy Sets. In: Petrosino, A. (ed.) WILF 2011. LNCS, vol. 6857, pp. 17–24. Springer, Heidelberg (2011)
37. Szmidt, E., Kacprzyk, J., Bujnowski, P.: How to measure the amount of knowledge conveyed by Atanassov's intuitionistic fuzzy sets. Information Sciences 257, 276–285 (2014)
38. Yuan, Y., Shaw, M.J.: Induction of fuzzy decision trees. Fuzzy Sets and Systems 69, 125–139 (1996)
39. Zadeh, L.A.: Fuzzy sets. Information and Control 8, 338–353 (1965)