An Approach to Intuitionistic Fuzzy Decision Trees

Paweł Bujnowski[1] **Eulalia Szmidt**[1,2] **and Janusz Kacprzyk**[1,2]

[1]Systems Research Institute, Polish Academy of Sciences, ul. Newelska 6, 01–447 Warsaw, Poland
[2]Warsaw School of Information Technology, ul. Newelska 6, 01-447 Warsaw, Poland

Abstract

An approach to construct a new classifier called an intuitionistic fuzzy decision tree is presented. Well known benchmark data is used to analyze the performance of the classifier. The results are compared to some other popular classification algorithms. Finally, the classifier behavior is verified while solving a real-world classification problem.

Keywords: Classification, decision tree, fuzzy decision tree, intuitionistic fuzzy decision tree

1. Introduction

One of the most popular classifiers with well known advantages are decision trees recursively partitioning a space of instances (observations). The ID3 algorithm [22] proposed by Quinlan is a source of many other approaches which have been developed along that line (cf. [27]).

Fuzzy decision trees (Janikow [17], Olaru et al. [21], Yuan and Shaw [41], Marsala [19], [20]), which are a generalization of classical (crisp) decision trees, turned out to be more stable, and more effective methods helping to extract knowledge while dealing with imperfect classification problems.

The expression power of using fuzzy sets in this context with respect to capturing and handling imprecision can significantly be enhanced by using various extensions of traditional concept of a fuzzy set. For instance, which is important for our work, the use of Atanassov's intuitionistic fuzzy sets [1], [2], [3] (A-IFSs for short) can provide an effective and efficient means for the representation and handling of imprecision in the setting of pro and con type statements and arguments, as well as hesitation. Needless to say that this kind of information and knowledge representation reflecting how humans proceed have been showed to be a powerful tool to solve many problems under imprecise information exemplified by machine learning and decision making, to name a few.

In this paper we present an approach to construct a new intuitionistic fuzzy decision tree classifier. The data is expressed by means of A-IFSs. Also the measures constructed for the A-IFSs are applied while making decisions how to split a node while expanding the tree. The intuitionistic fuzzy decision tree considered here is an extension of the fuzzy ID3 algorithm [7].

Well known benchmark data, and real-world data concerning an eye illness of premature born babies are applied to demonstrate the potential of the new algorithm. 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 [42]) given by

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

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

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

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

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

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. (An approach to the assigning memberships and non-memberships for A-IFSs from data is proposed by Szmidt and Baldwin [29]).

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

The hesitation margin turns out to be important while considering the distances (Szmidt and Kacprzyk [30], [31], [33], entropy (Szmidt and Kacprzyk [32], [34]), similarity (Szmidt and Kacprzyk [35]) for the A-IFSs, etc. i.e., the measures that play a crucial role in virtually all information processing tasks (Szmidt [28]).

The hesitation margin turns out to be relevant for applications – in image processing (cf. Bustince et al. [15], [14]), the classification of imbalanced and overlapping classes (cf. Szmidt and Kukier [36], [37], [38]), group decision making (e.g., [4]), genetic algorithms [24], negotiations, voting and other situations (cf. Szmidt and Kacprzyk papers).

3. Intuitionistic fuzzy decision tree

The intuitionistic fuzzy decision tree presented here was inspired by the soft decision tree introduced by Baldwin et al. [7] which, in turn, was an extension of the source *ID3* tree introduced by Quinlan [22]. The methods presented here make use of numeric attributes but they can also be applied to the nominal attributes (the algorithm is even simpler then). A-IFSs are used for data representation. Next, the new idea of deriving A-IFSs in each node was applied as potentially giving the most accurate results.

The process of a decision tree generation demands to point out the best attributes for splitting the nodes. 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 [32]) as a counterpart of "information gain" [22].

Below the most important components of the algorithm are described.

Fuzzy partitions of the attribute values (granulation)

Replacing a continuous domain with a discrete one, i.e., the idea of a universe partition (granulation), has been extended to fuzzy sets by Ruspini [25]. Here the idea was used 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. In other words, 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:

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

and for the j-th attribute A_j we have $o_{ij} \in A_j$, $i = 1, \ldots, n$, $j = 1, \ldots, m$.

It follows from (5) that the sum of the membership values for an observation o_{ij} is one (the sum results from only two neighboring fuzzy sets).

Remark. Symbol χ is used for the membership values for the purpose of granulation so to make a difference between membership values resulting from the attribute granulation (χ) and the membership values of the A-IFSs μ.

The following types of granulation are used:

– symmetric granulation (symmetric fuzzy partitions) with evenly spaced triangular fuzzy sets, and

– asymmetric granulation (asymmetric fuzzy partitions) with unevenly spaced triangular fuzzy sets such that each partition contains equal number of data points) [5, 25].

An example of both fuzzy partitioning (symmetric and asymmetric granulations) is shown in Fig. 1. The two kinds of partitioning are illustrated on attribute 2 of the "PIMA Diabetes" problem with 5 fuzzy sets (the "PIMA Diabetes" data was used as one of the benchmarks to assess intuitionistic fuzzy decision tree; summary results are in Table 3). Fuzzy partitioning (triangular fuzzy sets) points out how to assign nodes in a soft ID3 decision tree - cf. Fig. 2.

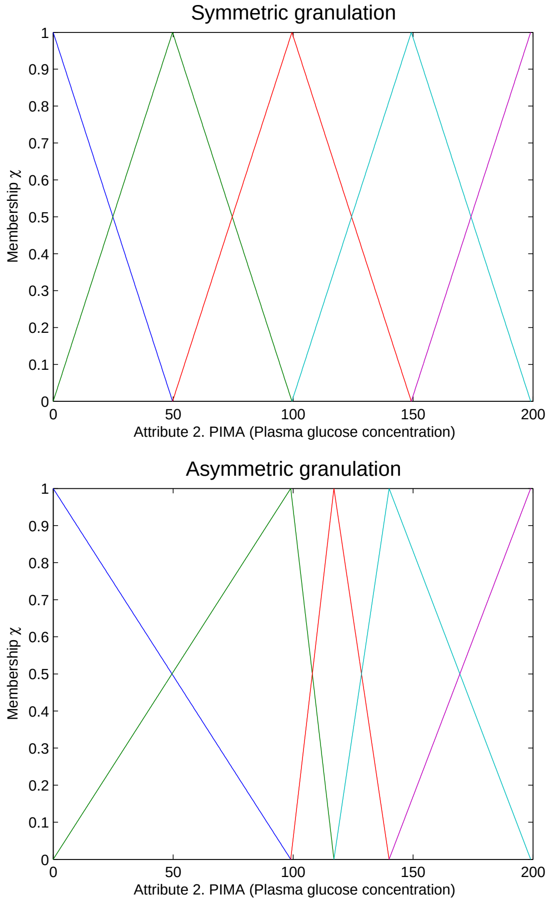

Figure 1: Example of symmetric fuzzy partitioning, and asymmetric fuzzy partitioning (on attribute 2 "Plasma glucose concentration" of benchmark "Pima Diabetes" with 5 fuzzy sets)

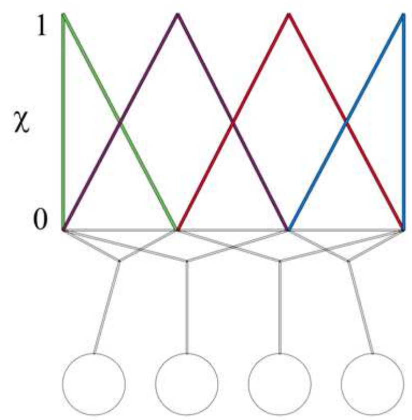

Figure 2: Constructing nodes in a soft ID3 tree resulting from a fuzzy partitioning

Now we will present a fuzzy generalization of *ID3* algorithm [7].

Fuzzy *ID3* algorithm

Consider the following database

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

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.

We assume that at the beginning the root contains all the instances, i.e., we apply top down approach of generating a fuzzy *ID3* decision tree from data. 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.

The rules can represent splitting the nodes in a decision tree. 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 triangular 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

$$B \equiv P_{\chi_{j_1}} \wedge \cdots \wedge P_{\chi_{j_N}} \quad (7)$$

where $P_{\chi_{j_r}}$ are triangular fuzzy sets, and its set of indices 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 triangular 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:

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

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 $\chi_{j_r(o_i)}$.

Consider $\{C_l, \ l = 1, \ldots, h\}$ 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

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

To split a node (starting from a root) it is necessary to evaluate the attributes' abilities 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:

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

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

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

It is assumed in (10) and (11) that A_s represents a rule from the root to the s-th child node.

The above formulas make it possible to generate the nodes in a fuzzy *ID3* tree [7].

Deriving A-IFSs from data

Making use of A-IFSs we will present now a generalization of the previously described soft *ID3* approach.

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

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

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)

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

Knowing the relative frequencies $p(C^+|A_i)$ and $p(C^-|A_i)$ (13) makes it possible to use the algorithm given in [6, 7] 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 [7] the fuzzy sets $Pos^+(A_s)$ and $Pos^-(A_s), \ s = 1, \ldots, p$ are tested by a classical entropy (10) - (11) to assess the attributes.

For the purpose of 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 [29]). Intuitionistic fuzzy model of the data in the child nodes $A_s, \ s = 1, \ldots p$ (due to [29]) is expressed by the following intuitionistic fuzzy terms

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

This way each child node s is described by the following A-IFS

$$< A_s, \mu(A_s), \nu(A_s), \pi(A_s) >, \ s = 1, \ldots, p \quad (15)$$

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

Characteristic of an instance o_i at node A_s can be expressed as well in terms of A-IFSs

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

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 $\{A_s,\ s=1,\dots,p\}$:

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

For the purpose of assessing and choosing the attributes while splitting the nodes in the intuitionistic fuzzy decision tree, either (15) or (16) may be used.

Selection of an attribute to split a node

Splitting a node into children nodes is the crucial step while expanding a tree – a crisp, fuzzy or intuitionistic fuzzy tree. 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 entropy [32].

Intuitionistic fuzzy entropy $E(x)$ of an intuitionistic fuzzy element $x \in A$ is [32]:

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

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 [31, 33]:

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

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

Other intuitionistic fuzzy measures may be used to evaluate the attributes (cf. [39], [40]), e.g.:

$$K(x)=1-0.5(E(x)+\pi_x), \quad (18)$$

where $\pi_x=1-\mu_x-\nu_x$ – (*hesitation margin, intuitionistic fuzzy index*) stands for the lack of knowledge concerning $x \in A$.

Intuitionistic fuzzy entropy $E(X)$ of an A-IFS with n elements: $X=\{x_1,\dots,x_n\}$ is [32]:

$$E(X)=\frac{1}{n}\sum_{i=1}^{n}E(x_i). \quad (19)$$

The same kind of calculations as (19) is performed also for the measure K.

We make use of the intuitionistic fuzzy representations (12)–(15) of the possible child nodes derived while testing attribute A to compute intuitionistic fuzzy entropy $E(A_s)$ (17) or the measure $K(A_s)$ (18) in a child node $A_s,\ s=1,\dots,p$.

Total intuitionistic fuzzy entropy of an attribute A is abbreviated $E(A)$ 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 $A_s,\ s=1,\dots,p$, with the weights reflecting supports (cardinalities) of the nodes:

$$E(A)=\frac{\sum_{s=1}^{p}w(A_s)E(A_s)}{\sum_{s=1}^{p}w(A_s)}. \quad (20)$$

Analogical formula is used for the total measure K.

An alternative way to (20) of calculating $E(A)$ is (or its counterpart $K(A)$) by applying a weighted intuitionistic fuzzy representation of each instance o_i (16) while partitioning an attribute A. Next, using (19), a total intuitionistic fuzzy entropy (or a total value of the measure K) is calculated for a chosen attribute. This method was applied in the numerical experiments (cf. Section 4).

An attribute with a minimal total intuitionistic fuzzy entropy (or a maximal value of the measure K) is selected for splitting a node.

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

Classification of the instances

Each leaf in a soft tree is described via a proportion of the classes considered. As a single instance usually belongs to several leaves, 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* being a sum of the products of the instance membership values at leafs and support for a class considered in these leafs [7]. Total support of the observation $o_i,\ i=1,\dots,n$, for a class C is:

$$supp(C|o_i)_{SUM}=\sum_{j=1}^{L}supp(C|T_j)\cdot\chi(T_j|o_i), \quad (21)$$

where: $\{T_j : j=1,\dots,L\}$ – a set of the leafs; L – the number of the leafs; $supp(C|T_j)$ – a support of the classes considered in the j-th leaf; $\chi(T_j|o_i)$ – 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

Behavior of the new intuitionistic fuzzy decision tree has been compared with other well known classification algorithms. The following measures were used in the process of the comparison:
– total proper identification of the instances belonging to the classes considered,
– the area under ROC curve [16].

Abilities of the intuitionistic fuzzy decision tree presented here were compared with the following classifiers:

- ***J48*** – implementation of the crisp tree proposed by Quinlan *C4.5* ([23]),
- ***LMT*** (*Logistic Model Tree*) – a hybrid tree with the logistic models at the leaves ([18]),
- ***NBTree*** – hybrid decision tree with the Bayes classifiers at the leaves,
- ***RandomForest*** – here consisting of 10 decision trees with nodes generated on the basis of a random set of attributes ([11]),
- ***MultilayerPerceptron*** – neural network,
- ***Logistic*** – logistic regression,

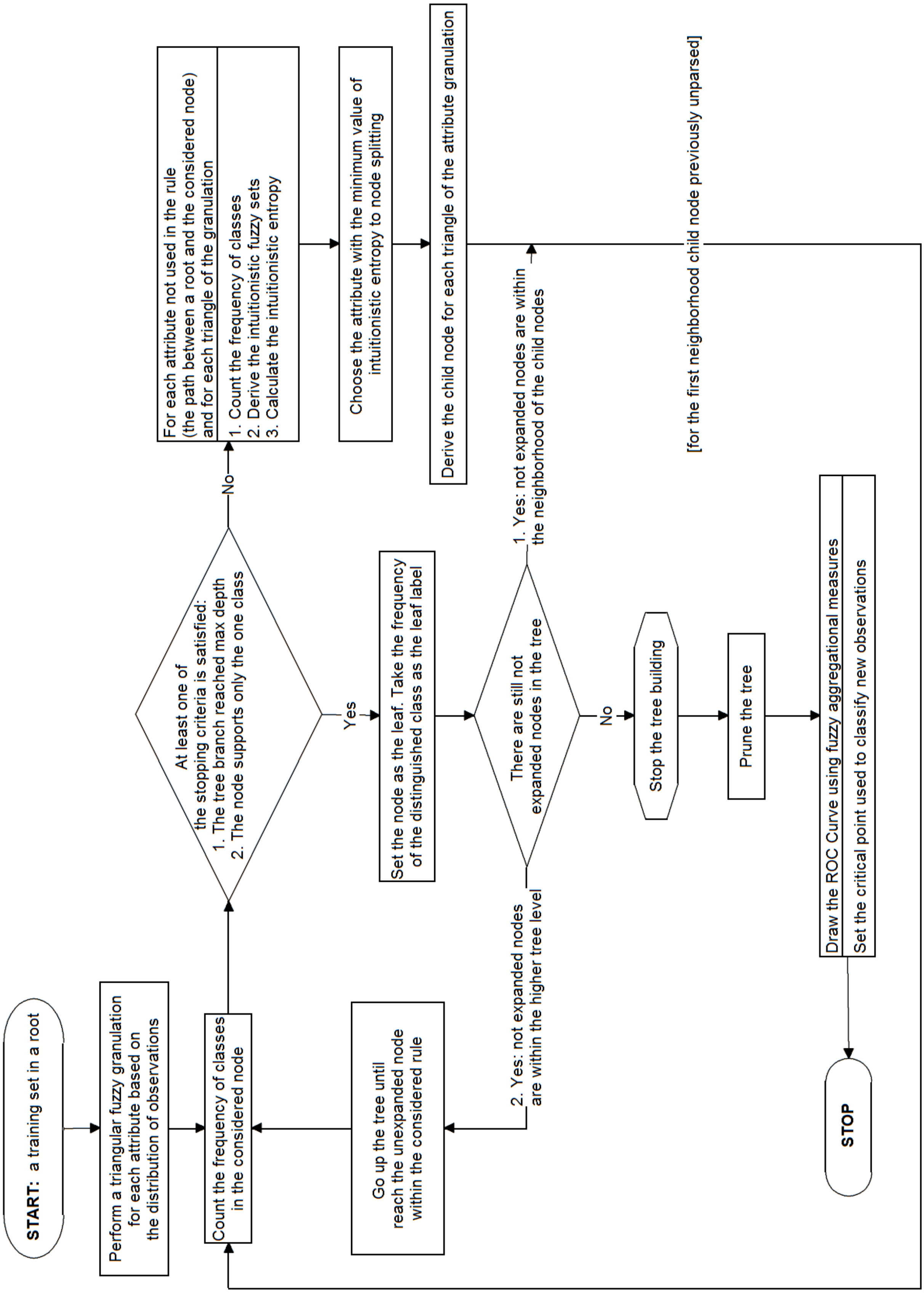

Figure 3: A flowchart representing a process of generating intuitionistic fuzzy tree

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

Algorithm	Classification accuracy ($\bar{x} \pm \sigma$) w % for all classes
$RandomForest$	77.05 ± 8.22
IFS tree (K, asym)	75.16 ± 6.21 $(*)$
pruned IFS tree (K, asym)	71.92 ± 6.30 $(-)$
SDT $(refitting)$	$71.09 \pm nd$ (nd)
$NBTree$	70.95 ± 9.95 $(-)$
SDT $(backfitting)$	$70.91 \pm nd$ (nd)
LMT	68.17 ± 9.91 $(-)$
$J48$ $(unpruned\, C4.5)$	68.07 ± 9.54 $(-)$
$J48$ $(pruned\, C4.5)$	67.61 ± 9.26 $(-)$
$MultilayerPerceptron$	65.96 ± 9.11 $(--)$
$LogisticModelTree$	63.92 ± 8.81 $(--)$

Table 2: “ROP” data – comparison of the intuitionistic fuzzy decision tree and other classifiers

Algorithm	Classification accuracy ($\bar{x} \pm \sigma$) w %	
	accuracy of both classes	AUC ROC
IFS tree (E, sym)	90.75 ± 2.52 $(*)$	90.73 ± 5.60 $(*)$
LMT	90.36 ± 2.53	90.16 ± 4.20
pruned IFS tree (E, sym)	91.50 ± 2.81 $(+)$	90.14 ± 4.72
$MultilayerPerceptron$	90.02 ± 3.10	86.82 ± 7.90 $(-)$
$RandomForest$	89.59 ± 3.08 $(-)$	86.48 ± 7.20 $(-)$
$Logistic$	90.68 ± 3.46	86.26 ± 9.02 $(-)$
$NBTree$	88.91 ± 2.70 $(-)$	81.58 ± 9.77 $(--)$
$J48$ $(unpruned\, C4.5)$	88.06 ± 3.14 $(-)$	74.12 ± 12.03 $(--)$
$J48$ $(pruned\, C4.5)$	88.72 ± 2.93 $(-)$	70.23 ± 14.76 $(--)$

Table 3: Ranking of the verified algorithms

Algorithm	Ranking of the results			
	Accuracy in respect to all classes		AUC ROC	
	$\bar{x} \pm \sigma$	$median$	$\bar{x} \pm \sigma$	$median$
LMT	3.0 ± 1.9	2.5	2.2 ± 1.1	2.0
IFS tree	2.8 ± 1.0	3.0	2.4 ± 1.3	3.0
$RandomForest$	4.4 ± 2.6	4.0	3.2 ± 1.9	3.0
$MultilayerPerceptron$	3.8 ± 2.6	3.5	4.0 ± 1.0	4.0
$Logistic$	4.6 ± 3.7	3	4.2 ± 2.6	5.0
$NBTree$	5.5 ± 1.6	5.0	5.2 ± 0.8	5.0
SDT $(backfitting)$	6.7 ± 1.2	6.5	nd	nd
SDT $(refitting)$	6.5 ± 2.2	7.0	nd	nd
$J48$ $(C4.5)$	6.9 ± 1.0	7.0	6.8 ± 0.4	7.0

- ***Soft Decision Trees (SDT)*** – proposed by Olaru and Wehenkel [21].

WEKA (http://www.cs.waikato.ac.nz/ml/weka/) was used to evaluate the above algorithms (excluding Soft Decision Trees (SDT) which results are presented in [21]).

To illustrate the results obtained by intuitionistic fuzzy decision tree we analyze first the results obtained for several benchmark data. Detailed results are given for benchmark data set “Glass” (http://archive.ics.-uci.edu/ml/datasets.html) containing 214 instances, 10 numerical attributes, 6 classes (4th class empty). “Glass” is not an easy data to analyze as the classes are imbalanced (e.g., class “building windows”contains 70 instances whereas “tableware” contains only 9 instances).

Next, analysis of the results is given for the data set “ROP” (retinopathy of prematurity) which is a real data set containing clinical data collected in a Polish hospital. ROP is a disease affecting eyes of the prematurely-born babies. Sometimes the disease can be mild but when not recognized early it may lead to blindness. It is the reason why an early classification of the babies with the disease is so important. Data set "ROP" contains 539 instances, 14 nominal attributes (e.g., sex, a kind of birth, different kinds of applied medical treatments like cardiac massage etc.) and 14 numerical attributes (e.g., general evaluation of a prematurely born using APGAR score, weight, gestational age, amount of given oxygen, etc.), 2 classes (with or without the disease).

Simple cross validation method is used 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 is calculated. *t-Student* test was used (Tables 1, 2) to compare an average accuracy of the new intuitionistic fuzzy decision tree with other classifiers. One minus in Tables 1, 2 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). By "nd" the cases are marked where no data is available.

Verifying the results for "Glass" benchmark data in respect to their accuracy (Table 1) we may notice that the intuitionistic fuzzy decision tree turned out a better classifier than other crisp and soft decision trees, even better than *Multilayer Perceptron* and *Logistic Model Tree*. Intuitionistic fuzzy decision tree placed itself on the second position being only a little worse than *Random Forest*.

Results for real data "ROP" (Table 2) show that the accuracy obtained by the intuitionistic fuzzy decision tree is the highest compared to the other verified classifiers (e.g., better than *Multilayer Perceptron*, *Random Forest*, *Logistic Model Tree*). It is also worth noticing that the standard deviation of the results is the lowest for the intuitionistic fuzzy decision tree which means that the classifier is most stable.

Final ranking of the tested algorithms (Table 3) in respect to all the examined data sets ("PIMA", "Ionosphere", "Sonar", "Wine", "Glass", "Iris" – http://archive.ics.uci.edu/ml/datasets.html, and two real data sets describing incomes, and the children illness "ROP") was done taking into account the average values with standard deviations ($\bar{x} \pm \sigma$), and the medians.

Table 3 presents results of the ranking in increasing order due to the median of the measure *AUC ROC*, and next, due to the median of the percentage of the proper identification of the classes. The intuitionistic fuzzy decision tree with its second position is only worse from a very effective hybrid tree *LMT*. It is worth stressing that a little worse than the intuitionistic fuzzy decision tree turned out *RandomForest* and *MutilayerPerceptron*.

Assessing the results of the verified algorithms (Table 3), besides the median, it is worth noticing as well the mean values and standard deviations of the ranking. Standard deviation for intuitionistic fuzzy tree is low in respect of both measures considered (because for each data set considered the tree did not obtained poor results). Again, it is worth emphasizing that the standard deviation of the accuracy is lower for intuitionistic fuzzy decision tree than for logistic regression (*Logistic*), random forest (*RandomForest*), and neural network (*MultilayerPerceptron*).

Last but no least, in many applications when transparency and comprehensibility to the human being is relevant, the proposed classifier, as a tree type classifier, can be a properer, if not the best choice.

5. Conclusions

We have presented and tested an extension of the fuzzy ID3 decision tree algorithm, namely, a new intuitionistic fuzzy decision tree. The new classifier was tested on well known benchmark examples and real-world data examples giving very encouraging results.

Acknowledgment

Partially supported by the National Science Centre Grant UMO-2012/05/B/ST6/03068.

.

References

[1] Atanassov K. (1983): *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. (1999): *Intuitionistic Fuzzy Sets: Theory and Applications.* Springer.

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

[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., Karale S. B. (2003): *Asymmetric Triangular Fuzzy Sets for Classification Models.* in V. Palade, R. J. Howlett, L. C. Jain (Eds): KES 2003, LNAI 2773, Springer-Verlag Berlin Heildelberg, 364–370.

[6] Baldwin J. F., Lawry J., Martin T. P. (1996): *A mass assignment theory of the probability of fuzzy events.* Fuzzy Sets and Systems, 83, 353–367.

[7] Baldwin J. F., Lawry J., Martin T. P. (1997): *Mass Assignment Fuzzy ID3 with Applications.*. Unicom Workshop on Fuzzy Logic Applications and Future Directions, London.

[8] Bartczuk L and Rutkowska D. (2006): *A New Version of the Fuzzy-ID3 Algorithm*. In: Artificial Intelligence and Soft Computing - 8th Int. Conf. ICAISC 2006, L. Rutkowski et al. (Eds.), Lecture Notes in Computer Science, Springer, 1060–1070.

[9] Benbrahim H., Bensaid A. (2000): *A comparative study of pruned decision trees and fuzzy decision trees.* NAFIPS 2000, 227–231.

[10] Bezdek J. C. (1981): *Pattern Recognition with Fuzzy Objective Function Algorithms.* Kluwer Academic Publishers Norwell, MA, USA.

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

[12] Breiman L., Friedman J. H., Olsen R. A., Stone C. J. (1984): *Classification and Regression Trees.* Wadsworth, Belmont, CA.

[13] Bujnowski P. (2013): Using intuitionistic fuzzy sets for constructing decision trees in classification tasks. PhD dissertation (in Polish), IBS PAN, Warsaw.

[14] Bustince H., Mohedano V., Barrenechea E., and Pagola M. (2005): *Image thresholding using intuitionistic fuzzy sets.* In: Issues in the Representation and Processing of Uncertain and Imprecise Information. Fuzzy Sets, Intuitionistic Fuzzy Sets, Generalized Nets, and Related Topics. (Eds. Atanassov K., Kacprzyk J., Krawczak M., Szmidt E.), EXIT, Warsaw.

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

[16] 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.

[17] Janikow C. Z. (1998): *Fuzzy Decision Trees: Issues and Methods.* IEEE Transactions on Systems, Man, and Cybernetics, 28, 1, 1–14.

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

[19] Marsala Ch. (2000): *Fuzzy decision trees to help flexible querying.* Kybernetika, 36 (6), 689–705.

[20] Marsala Ch. and Bouchon-Meunier B. (1999): *An adaptable system to construct fuzzy decision tree.* NAFIPS'99, 223–227.

[21] Olaru C., Wehenkel L. (2003): *A complete fuzzy decision tree technique.* Fuzzy Sets and Systems, Elsevier, 221–254.

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

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

[24] 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.

[25] Ruspini E. H. (1969): *A New Approach to Clustering.* Information and Control, 15, 22–32.

[26] Rutkowski L. (2009): *Artificial intelligence methods and techniques.* PWN, Warszawa, 237–307. (in Polish).

[27] Safavian S.R. and Landgrebe D. (1991): *A survey of decision tree classifier methodology.* IEEE Trans. Systems Man Cybernet. 21, 660–674.

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

[29] Szmidt E. and Baldwin J. (2006): *Intuitionistic Fuzzy Set Functions, Mass Assignment Theory, Possibility Theory and Histograms.* 2006 IEEE WCCI, 237–243.

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

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

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

[33] Szmidt E., Kacprzyk J. (2006): *Distances Between Intuitionistic Fuzzy Sets: Straightforward Approaches may not work.* 3rd International IEEE Conference Intelligent Systems IEEE IS'06, London, 716–721.

[34] Szmidt E. and 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

[35] 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

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

[37] 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.

[38] 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. Series: Studies in Computational Intelligence.

[39] Szmidt E., Kacprzyk J., Bujnowski P. (2011): *Measuring the Amount of Knowledge for Atanassov's Intuitionistic Fuzzy Sets.* Fuzzy Logic and Applications, Lecture Notes in Artificial Intelligence, Vol. 6857, 2011, 17–24.

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

[41] Yuan Y., Shaw M. J. (1996): *Induction of fuzzy decision trees.* Fuzzy Sets and Systems 69, 125–139.

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