Proceedings of the Federated Conference on
Computer Science and Information Systems pp. 213–219
ISBN 978-83-60810-51-4

Fuzzy Decision Trees in Medical Decision Making Support System

Vitaly Levashenko, Elena Zaitseva
University of Zilina,
Univerzitna 8215/1,
01026, Zilina, Slovakia
Email: {vitally.levashenko, elena.zaitseva}@fri.uniza.sk

***Abstract*—Decision Making Support System based on Fuzzy Logic is considered in this paper for oncology disease diagnosis. The decision making procedure corresponds to the recognition (classification) of the new case by analyzing a set of instances (already solved cases) for which classes are known. Ontology (solved cases) is defined as Fuzzy Classification Rules that are formed by different Fuzzy Decision Trees. Three types of Fuzzy Decision Trees (Non-ordered, ordered and Stable) are considered in the paper. Induction of these Fuzzy Decision Trees is based on Cumulative Information Estimates. The proposed approach is implemented based on medical problem benchmark with real clinical data for breast cancer diagnosis.**

I. Introduction

DATA mining is a process of extracting implicit, potential, novel, useful and intelligible patterns from mass data of data sets, databases or data warehouse, etc. The technologies of classification, estimation, prediction, affinity grouping, association rules, clustering, description and visualization are covered in data mining, which is widely used in the fields of medicine [1]. Decisions play an important role in medicine, especially in medical diagnostic processes. *Decision Making Support Systems* (DMSS) helping physicians are becoming a very important part in medical decision making, particularly in those situations where decision must be made effectively and reliably. Since conceptual simple decision making models with the possibility of automatic learning should be considered for performing such tasks, decision trees are a very suitable candidate. They have been already successfully used for many decision making purposes [1]–[3].

A decision tree is a graphic model of a decision process, and it is usually used as a decision support tool or classifier. A decision trees is one of the best ways to analyze a decision, as it is visualized and simple to understand and interpret. Its possible consequence includes chance event outcomes, resource costs or utility.

We propose to investigate DMSS based on Fuzzy Logic and *Fuzzy Decision Trees* (FDT), as an efficient alternative to crisp classifiers that are applied independently. An important aspect of this model is cooperation of Fuzzy Logic and decision trees. This cooperation tries to soften the accuracy/interpretability tradeoff.

Induction of FDT is a useful technique to find patterns in data in the presence of imprecision, either because data are fuzzy in nature or because we must improve its semantics. We have proposed the technique for induction of new type of fuzzy decision tree – ordered FDT, which is simple to understand and apply. The use of cumulative information estimations allows precisely estimating mutual influence of attributes. These evaluations are used to analyze group of training instances.

Many FDT induction algorithms have been considered and introduced in [4]. There are different medical applications of FDT for building of rules for classification [2], [3], [5]. Authors in papers [2], [5] use fuzzy ID3 algorithm, that doesn't allow FDT building with parallel structure. FDT in [3] is satisfactory for completely specified initial data.

In [6] the ordered FDT have been proposed that permit to find a sequence of rules, which analyze input attributes in order that is both cost effective and guarantees a desired level of accuracy. Every node of one level of such FDT associates with similar attribute.

For these purposes, a technique to compute cumulative information estimates of fuzzy sets have been used [7]. The application of such estimations allows inducing minimum cost decision trees based on new criterions of optimality.

In this paper we develop application of cumulative information estimations and propose new type of FDT – stable FDT that can be used for incompletely specified initial data. Therefore three types of FDT are considered in this paper: non-ordered FDT (is inducted by fuzzy ID3 algorithm), ordered FDT and stable FDT. The non-ordered FDT is inducted by fuzzy ID3 algorithm. The ordered and stable FDT are build based on cumulative information estimations. In this paper we consider application of these

This work was supported by grant of Slovak Electrical Society Section with the Department of Telecommunications University of Zilina, Slovakia

types FDT for medical application, particularly for prognostic of breast cancer.

The paper is organized as follows. Section II contains brief information about used DMSS and representation of fuzzy data. Section III and shows three types of FDT induction with a simple example. Section IV contains description of our Fuzzy classification rules. Section V demonstrates the results of tests by benchmark for breast cancer diagnosis.

II. Decision Making Support System based on Fuzzy Logic

A. Decision Making Support System

The huge amount of medical data and the different sources of medical information make the task of decision making difficult and complex. DMSS are systems used in medicine for the tasks of diagnosis, prognosis, treatment planning and decision support [8], [9]. There are a lot of such systems that have been implemented for specific areas in medicine or diseases. Some of these systems have similar conception and based on identical mathematical background. Therefore we consider a conception of the DMSS and its application for prognosis of some diseases.

There are some conceptions of the DMSS structure [9, 10]. We use conception with comparison of a new case with previous cases and selection most similar as decision (Fig.1). Thus the classification is principal problem of this conception based on special rules that agrees with Block of Compare new case and ontology. The mathematical background of this block is Fuzzy Classification rules that are formed by FDT. The block for Preparation of initial data implements transformation of the input data to the fuzzy data. This procedure is named as fuzzification. The result presentation is interpretation of the decision by the de-fuzzification procedure.

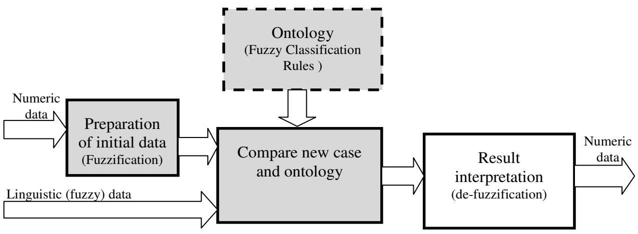

Fig. 1 Decision Making Support System

The decision making procedure corresponds to the recognition (classification) of the new case and is the process of moving from concrete examples to general models, where the goal is to learn how to classify objects by analyzing a set of instances (already solved cases) whose classes are known. Instances are typically represented as attribute-value vectors. One of possible solutions for such classification is implemented by Decision Trees [10]. A decision tree is formalism for expressing such mappings and consists of tests of attribute nodes linked to two or more sub-trees and leafs or decision nodes labeled with a class which means the decision. A test node computes some outcome based on the attribute values of an instance, where each possible outcome is associated with one of the sub-trees. An instance is classified by starting at the root node of the tree. If this node is a test, the outcome for the instance is determined and the process continues using the appropriate sub-tree. When a leaf is eventually encountered, its label gives the predicted class of the instance. The FDT is one of possible types of decision trees that permits to operate by fuzzy data (attributes).

The process of construction of FDT is based on the use of a fuzzy partition for each numerical attribute. An automatic method of construction of such a partition from a set of precise values could be used in order to obtain automatically a set of fuzzy values for each numerical attribute. Fuzzy data are used in situations that are especially difficult or ambiguous, and unsolvable by other types of logic. Fuzzification transforms precise input into corresponding fuzzy input [11]. The interpretability of a fuzzy system – especially if applied in data analysis – is one of its key advantages. Therefore the considered DMSS (Fig.1) is implemented based on the fuzzy classification rules.

B. Fuzzy Logic

Fuzzy logic is a popular approach to capture vagueness of information [12]. The basic idea is to use instead the “crisp” 1 and 0 values the values of the interval [0,1] indicating a degree of truth or confidence.

A fuzzy set F with respect to a universe U is characterized by a membership function μ_F: U $\rightarrow$ [0,1], which assign a F-membership degree, $\mu_F(u)$, to each element u in U. $\mu_F(u)$ gives an estimation that u belongs to the fuzzy set F [13].

For example, consider attribute A_i that is Age. This attribute has 3 fuzzy partitions $A_{i,1}$ (young), $A_{i,2}$ (adult), $A_{i,3}$ (old) (with range [0,1]) as it is depicted in Fig.2. The real value $u \in U$ of this attribute A_i is interpreted as: $\mu_{young}(u)=0.7$, $\mu_{adult}(u)=0.3$, and $\mu_{old}(u) = 0$ in terms of fuzzy logic.

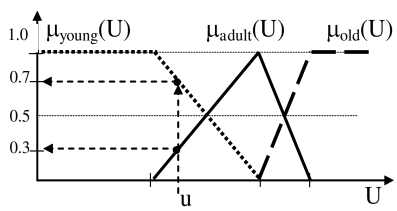

Fig. 2 Fuzzy membership functions of an attribute A_i (Age)

Thus, the fuzzification of the initial data is performed by analyzing the corresponding values of a membership function. Here, each attribute value can be seen as likelihood estimate. In this paper we analyze a particular case when the sum of membership values of all partitions equals to 1. For these purposes, we use one of the algorithms to transform from numeric to triangular fuzzy data, presented in [11].

A typical classification problem can be described as follows [14]. A universe of objects U={u} is described by N training examples and n input attributes $\mathbf{A}=\{A_1,...,A_n\}$. Each attribute A_i ($1 \le i \le n$) measures some feature presented by a group of discrete linguistic terms. We assume that each group is a set of m_i ($m_i \ge 2$) values of fuzzy subsets $\{A_{i,1},\ldots,A_{i,j},\ldots,A_{i,m_i}\}$.

The cost of an attribute A_i denoted as Cost(A_i) is an integrated measure that accounts financial and temporal costs required to define the value of the A_i for an instance. We assume that each object u in the universe is classified by a set of classes $\{B_1,...,B_{m_b}\}$. This set describes the class attribute B. The class attribute B has to determine by values of attributes A_i with minimal costs Cost(A_i).

Let us consider the simplified example for breast cancer diagnosis (Table I). In this example we use only four inputs attributes [15]: A_1 (Gynecological history), A_2 (Tumor), A_3 (Heredity), and A_4 (Age) and let each instance be connected to one class attribute B (Breast Cancer Possibility). Let each attribute has the values: $A_1=\{A_{1,1}, A_{1,2}, A_{1,3}\}$, $A_2=\{A_{2,1}, A_{2,2}, A_{2,3}\}$, $A_3=\{A_{3,1}, A_{3,2}\}$, $A_4 = \{A_{4,1}, A_{4,2}\}$ and $B = \{B_1, B_2, B_3\}$. Let the instances be the ones presented in Table I. Let the costs of attributes be the ones on the last row of the Table II.

TABLE I.
ATTRIBUTES VALUES

Attribute	Attribute Values	Description of Attribute Values
A_1	$A_{1,1}$	Gynecological history with high risk
	$A_{1,2}$	Gynecological history with medium risk
	$A_{1,3}$	Gynecological history with low risk
A_2	$A_{2,1}$	Yes and confirmed by medical examination
	$A_{2,2}$	Yes and non-confirmed
	$A_{2,3}$	No
A_3	$A_{3,1}$	Yes
	$A_{3,2}$	No
A_4	$A_{4,1}$	Younger than 40 years
	$A_{4,2}$	Over 40 years
B	B_1	High Possibility of Breast Cancer
	B_2	Medium Possibility of Breast Cancer
	B_3	Low Possibility of Breast Cancer

III. FUZZY DECISION TREES INDUCTION

There are different approaches to induct FDT [16], [17]–[19]. The principal goal of these approaches for FDT induction is selection of expanded attributes and determination of the leaf node. The key points of approaches for induction of FDT are (a) a heuristic for selecting expanded attributes and (b) a rule for transformation of nodes into leaves. An expanded attribute is an attribute that according to the values of the attribute tree expands the node considered.

In paper [7] new cumulative information estimates have been proposed. The cumulative information estimates allow defining criterion of expanded attributes selection to induct FDT with different properties. Non-ordered, Ordered and Stable FDT will be considered below. These FDT are different by selection criterion of expanded attributes. This selection criterion is defined as different type of cumulative mutual information **I**(B; **A**), where B and **A** are output and input attributes (or its values).

The selection criterion of expanded attributes A_{i_q} for induction of Non-ordered FDT is defined as [6, 7]:

$$\frac{\mathbf{I}(B; A_{i_1,j_1},\ldots,A_{i_{q-1},j_{q-1}}, A_{i_q})}{Cost(A_{i_q})} \to \max , \quad (1)$$

where $A_{i_1,j_1},\ldots,A_{i_{q-1},j_{q-1}}$ are values of input attributes $A_{i_1},\ldots,A_{i_{q-1}}$ of path from root node to examined attribute; A_{i_q} is the attribute that isn't in this path.

In (1) the Cost(A_i) is an integrated measure that accounts

TABLE II.
A TRAINING SET

No	A_1			A_2			A_3		A_4		B		
	$A_{1,1}$	$A_{1,2}$	$A_{1,3}$	$A_{2,1}$	$A_{2,2}$	$A_{2,3}$	$A_{3,1}$	$A_{3,2}$	$A_{4,1}$	$A_{4,2}$	B_1	B_2	B_3
1.	0.9	0.1	0.0	1.0	0.0	0.0	0.8	0.2	0.4	0.6	0.0	0.8	0.2
2.	0.8	0.2	0.0	0.6	0.4	0.0	0.0	1.0	0.0	1.0	0.6	0.4	0.0
3.	0.0	0.7	0.3	0.8	0.2	0.0	0.1	0.9	0.2	0.8	0.3	0.6	0.1
4.	0.2	0.7	0.1	0.3	0.7	0.0	0.2	0.8	0.3	0.7	0.9	0.1	0.0
5.	0.0	0.1	0.9	0.7	0.3	0.0	0.5	0.5	0.5	0.5	0.0	0.0	1.0
6.	0.0	0.7	0.3	0.0	0.3	0.7	0.7	0.3	0.4	0.6	0.2	0.0	0.8
7.	0.0	0.3	0.7	0. 0	0.0	1.0	0.0	1.0	0.1	0.9	0.0	0.0	1.0
8.	0.0	1.0	0.0	0.0	0.2	0.8	0.2	0.8	0.0	1.0	0.7	0.0	0.3
9.	1.0	0.0	0.0	1.0	0.0	0.0	0.6	0.4	0.7	0.3	0.2	0.8	0.0
10.	0.9	0.1	0.0	0.0	0.3	0.7	0.0	1.0	0.9	0.1	0.0	0.3	0.7
11.	0.7	0.3	0.0	1.0	0.0	0.0	1.0	0.0	0.2	0.8	0.3	0.7	0.0
12.	0.2	0.6	0.2	0.0	1.0	0.0	0.3	0.7	0.3	0.7	0.7	0.2	0.1
13.	0.9	0.1	0.0	0.2	0.8	0.0	0.1	0.9	1.0	0.0	0.0	0.0	1.0
14.	0.0	0.9	0.1	0.0	0.9	0.1	0.1	0.9	0.7	0.3	0.0	0.0	1.0
15.	0.0	0.0	1.0	0.0	0.0	1.0	1.0	0.0	0.8	0.2	0.0	0.0	1.0
16.	1.0	0.0	0.0	0.5	0.5	0.0	0.0	1.0	0.0	1.0	0.5	0.5	0.0
Cost(A_i)	2,5			1,7			2,0		1,8				

financial and temporal costs required to define the value of the A_i for an instance and this value is defined a priori.

Maximum value of cumulative mutual information (1) allows to select expanded attributes between attributes A_{i_q}.

There are two tuning parameters α and β used in the algorithm [6, 7]. Expanding a tree branch is stopped when either the frequency f of the branch is below α or when more than β percent of instances left in the branch has the same class label. Thus these values are key parameters needed in the fifth step of the algorithm deciding have we already approached to leaf node or should we need to continue expanding the branch.

The Non-ordered FDT inducted according (1) for data in Table I (β=0,75 and α=0,16) is in Fig.3. The cumulative conditional entropy $\mathbf{H}(B|\mathbf{A})$ is used in induction of this FDT. In paper [7] this entropy is presented in detail.

According to this FDT the attribute A_2 has maximum value of information estimation (1) therefore the analysis of new patient findings start of this attribute. The value of this attribute $A_{2,3}$ (frequency f = 0.269) causes output attribute B as B_3 with value 78.8% (Low Possibility of Breast Cancer). This value is more than β = 0.75 therefore the analysis of new patient findings is finished. The values of attribute A_4 is considered if the attribute A_2 has value $A_{2,2}$. The output attribute B is B_3 with value 75.4% if A_4 is $A_{4,1}$. Other value of attribute A_4 provides estimation of the attribute A_1. The output attribute B is B_1 with value 57.9% or 55.1% (High Possibility of Breast Cancer) if A_1 is $A_{1,1}$ or $A_{1,2}$ accordingly; and B is B_3 with value 49.5% if A_1 is $A_{1,3}$. Here the output attribute values B_1 and B_3 are less than β = 0.75, but the frequency value f is less than α = 0.16 therefore the analysis is finished. The value $A_{2,1}$ of the attribute A_2 causes the estimation of the attribute A_1 in the first and the attribute A_4 in the second. The analysis of new patient findings is finished if A_1 is $A_{1,3}$ with output value B_3. The output attribute B is B_2 (Medium Possibility of Breast Cancer) if attribute A_1 is $A_{1,2}$. The analysis is stopped because the frequency f is less than α for $A_{1,2}$ and $A_{1,2}$. The attribute values $A_{4,1}$ and $A_{4,2}$ causes output attribute B_2 with value 67.8% and 62.3% accordingly. The frequency values f = 0.088 and f = 0.151 are below than value of parameter α therefore the analysis is finished. The attribute A_3 for parameters α = 0.16 and β = 0.75 doesn't influence to analysis of new patient findings.

The induction of an Ordered FDT is less complex, when it does not require calculation of information estimates for each branch of the tree. Choosing an expanded attribute A_{i_q} according (2) is sufficient enough to maximize the increment of information of the attribute at minimum costs. Usage of cumulative information estimates forms a criterion for ordered FDT induction:

$$\frac{\mathbf{I}(B; A_{i_1}, \dots, A_{i_{q-1}}, A_{i_q})}{Cost(A_{i_q})} \to \max . \quad (2)$$

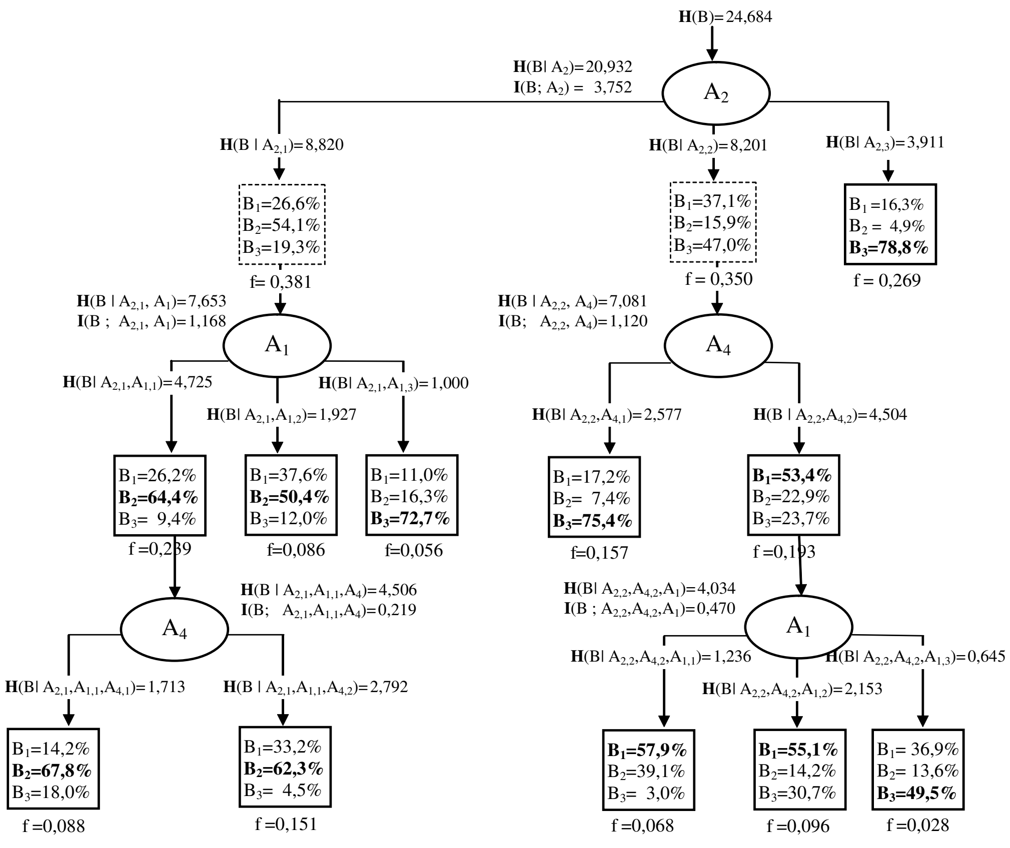

Fig. 3 Non-ordered FDT inducted for the data in Table I

Algorithm for Ordered FDT induced has been presented in [6]. The Ordered FDT induced from the data presented in Table I can be seen in Fig. 4. The Ordered FDT includes equal attributes in every level. Therefore the predefined order of attributes allows parallel calculation by this tree.

Modification of Ordered FDT is Stable FDT. Stable FDT are best for classification problem if one or more attributes are absent. The criterion for induction of stable FDT based on of cumulative information estimates is:

$$\frac{\mathbf{I}(A_{i_q};B,A_{i_1},\ldots,A_{i_{q-1}})}{\mathrm{Cost}(A_{i_q})}\to \max\ , \qquad (3)$$

The Stable FDT is in Fig. 5 for the data presented in Table I. The more important (by information estimates) attributes is located in root or near root of the tree. Therefore the stable result will be obtained but the cost $\mathrm{Cost}(A_i)$ has to increase.

IV. Usage of FDT for classification

In the all type of FDT, each non-leaf node is associated with an attribute $A_i \in A$. When A_i is associated with a non-leaf node, the node has m_i outgoing branches. The j-th branch of the node is associated with value $A_{i,j}$. The class attribute B has m_b possible values $B_1,\ldots, B_{j_b},\ldots, B_{m_b}$. Let the FDT have R leaves $L = \{l_1, \ldots, l_r, \ldots, l_R\}$. There is also a vector of values $\mathbf{B}^r = [B_1{}^r;\ldots, B_{j_b}{}^r;\ldots, B_{m_b}{}^r]$ for each r-th leaf l and each j_b-th class B_{j_b}. Each value $F_{j_b}{}^r$ means the certainty degree of the class B_{j_b} attached to the leaf node l_r.

In fuzzy cases, a new instance e may be classified into different classes with different degrees. Then, each leaf $l_r \in L$ corresponds to one (r-th) classification rule. The condition part of the classification rule is a group of conditions presented in the form "attribute is attribute is value" and those conditions are connected with and-operator. These attributes are associated with the nodes in the path from the root to the leaf l_r. The attribute's values are the values associated with the respective outgoing branches of the nodes in the path. The conclusions of the r-th rule are the values of class attribute B with their truthfulness vector $\mathbf{B}^r$ values.

Let's consider the path for FDT (Fig.4) $P_r(e)=\{[A_{i_1,j_1}(e)]^r,\ldots,[A_{i_s,j_s}(e)]^r,\ldots,[A_{i_S,j_S}(e)]^r\}$ from the root to the r-th leaf . This path $P_r(e)$ consists of S nodes which are associated with attributes $A_{i_1},\ldots, A_{i_s},\ldots, A_{i_S}$ and respectively their S outgoing branches associated with the values $A_{i_1,j_1},\ldots, A_{i_s,j_s},\ldots, A_{i_S,j_S}$. Then the r-th rule has the following form:

IF (A_{i_1} is A_{i_1,j_1}) and … and (A_{i_S} is A_{i_S,j_S}) THEN B
(with truthfulness $\mathbf{B}^r$).

Our approach uses several classification rules for classification of a new instance e. That's why, there may be several paths whose all outgoing node's branches are associated with values $A_{i_S,j_S}(e)$ greater than 0. Each path $P_r(e)$ form the root to the leaf node l_r corresponds to one r-th

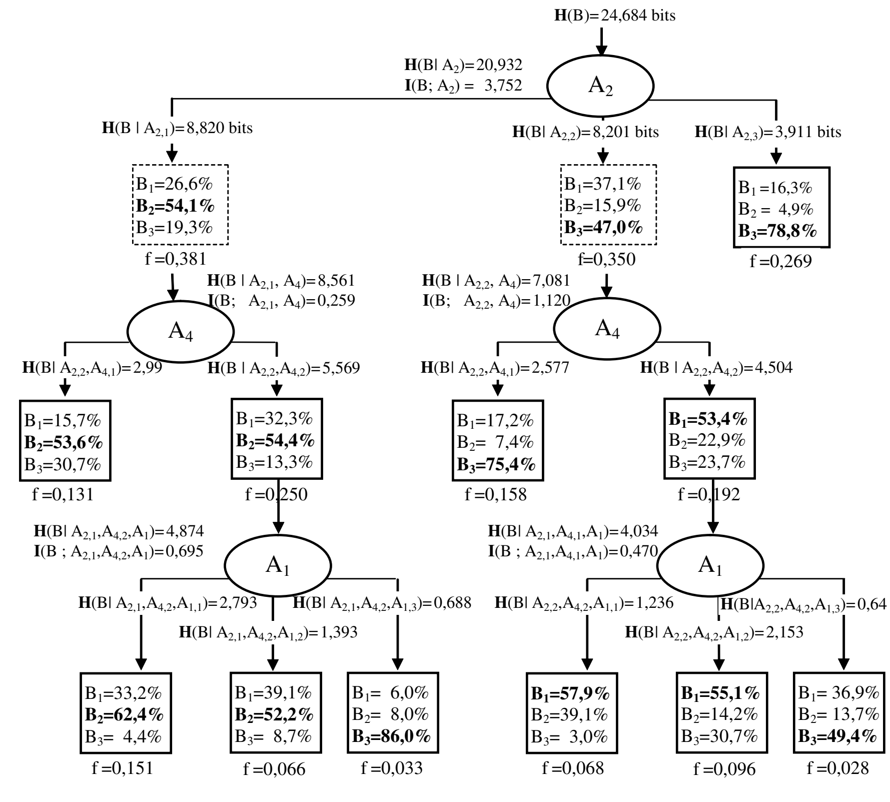

Fig. 4 Ordered FDT inducted for the data in Table I

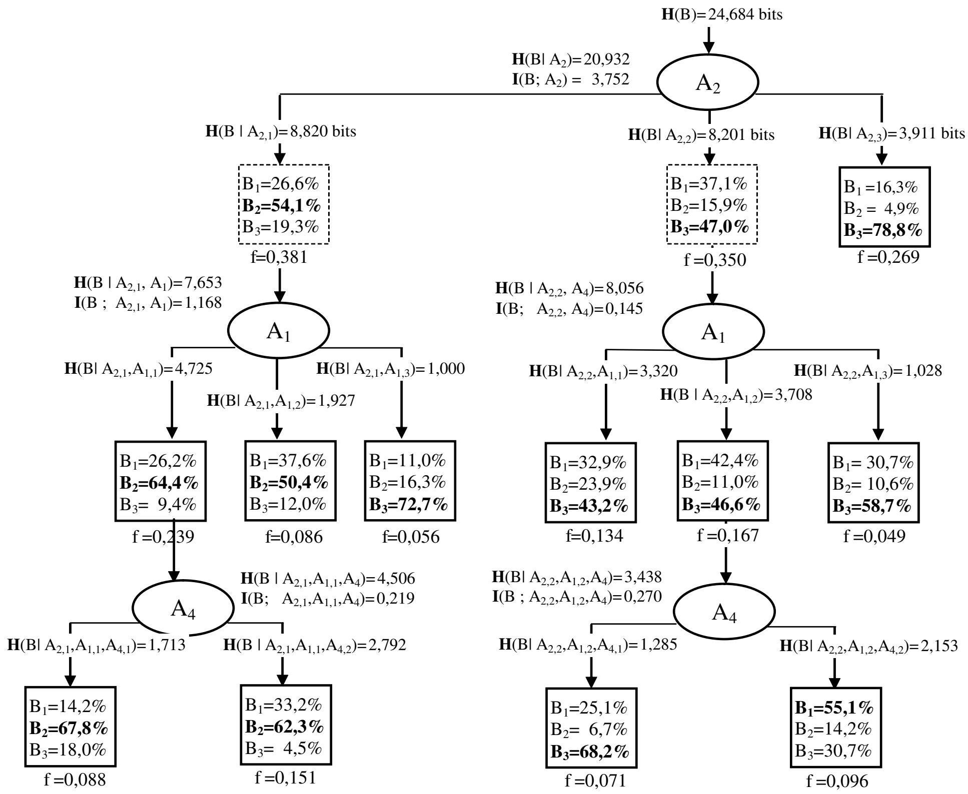

Fig. 5 Stable FDT inducted for the data in Table I

classification rule. In this case each r-th classification rule should be included in the final classification with a certain weight $W_r(e)$. The weight for instance e and the r-th rule is given by the rule $W_r(e)=\prod_{s=1}^{S}[A_{i_s,j_s}(e)]^r$, where $[A_{i_S j_S}(e)]^r$ is the value of the attribute A_{i_S}, for the new instance e. The weight $W_r(e)$ is equal 0 if there is a attribute's value $A_{i_S j_S}$ whose membership function equals 0. Values of class attribute B for the new instance e are:

$$\mu_B(e) = \sum_{r=1}^{R} W_r(e) \times \mathbf{B}^r\,, \qquad (4)$$

where $\mathbf{B}^r$ is the truthfulness of the r-th rule.

Below the transformation process of the FDT into fuzzy rules and these rules are used for classification are described by example for the Ordered FDT in Fig.4.

The FDT in Fig.4 has R=9 leaves. Let a new instance e have following attribute values: $A_1=\{A_{1,1};A_{1,2};A_{1,3}\} = \{0.9; 0.1; 0.0\}$, $A_2 =\{A_{2,1};A_{2,2};A_{2,3}\}= \{1.0; 0.0; 0.0\}$, $A_3=\{A_{3,1}; A_{3,2}\}=\{0.8; 0.2\}$ and $A_4=\{A_{4,1};A_{4,2}\} =\{0.4; 0.6\}$. Our goal is to determine values of class attribute B for this new instance e.

Let's form 9 classification rules for the FDT leaves.

r=1: IF A_2 is $A_{2,1}$ and A_4 is $A_{4,1}$ THEN B
truthfulness $\mathbf{B}^1$ = [0.157; 0.536; 0.307];

r=2: IF A_2 is $A_{2,1}$ and A_4 is $A_{4,2}$ and A_1 is $A_{1,1}$ THEN B
truthfulness $\mathbf{B}^2$ = [0.332; 0.624; 0.044];

r=3: IF A_2 is $A_{2,1}$ and A_4 is $A_{4,2}$ and A_1 is $A_{1,2}$ THEN B
truthfulness $\mathbf{B}^3$ = [0.391; 0.522; 0.087];

…

r=9: IF A_2 is $A_{2,3}$ THEN B
truthfulness $\mathbf{B}^9$ = [0.163; 0.049; 0.788].

The weights $W_r(e)$ (r=1,…,9) are: $W_1(e)=1.0\times0.4=0.4$, $W_2(e)=1.0\times0.6\times0.9=0.54$, $W_3(e)=1.0\times0.6\times0.1= 0.06$, and all the other $W_r(e)$ are equal 0.

We obtain by use of (4) for this FDT:

$\mu_{B_1}(e)=0.157\times0.4+0.332\times0.54+0.391\times0.06 = 0.265$.

Similarly,

$\mu_{B_2}(e) = 0.536\times0.4 + 0.624\times0.54 + 0.522\times0.06 = 0.583$;

$\mu_{B_3}(e) = 0.307\times0.4 + 0.044\times0.54 + 0.087\times0.06 = 0.152$.

The values of class attribute B = $\{B_1, B_2, B_3\}$ = {0.265; 0.583; 0.152} for the new instance e. The maximum value has $\mu_{B_2}(e)$. And so, if classification only into one class is needed, instance e is classified into class B_2.

V. Illustrative Application and Experimental Results

Analysis and application of presented types of FDT in previous section are considered in this section. We used data for diagnosis of breast cancer to form classification rules based on non-ordered, ordered and stabile FDT. The experiments have been carried out on Machine Learning benchmarks (dataset) each of which has [20], [21]: breast-cancer-wisconsin and breast-cancer. We had divided initial dataset into 2 parts. The first part (70% from initial dataset) was used for building classification models. The second part (30% from initial dataset) was used for verification of the classification models. This process was repeated 1000 times, and average estimations were produced.

A fragment of our results is shown in Table III. Columns [Total sets], [Number of attributes] and [Number of classes] describes dataset. The column labeled [Errors] gives the count of error classification. It is calculated as the ratio of the number of misclassification combinations to the total number of combinations. The results in columns Non-oFDT, oFDT, sFDT are according to fuzzy classification rules that have been formed based on Non-ordered, Ordered and Stable FDT. Note, the Stable FDT has been inducted in condition where one or two attributes were absent.

TABLE III.
RESULTS ON THE UCI MACHINE LEARNING BENCHMARK SET

Dataset	Total sets	Number of input attributes	Number of classes	Errors		
				Non-oFDT	oFDT	sFDT
breast-cancer-wisconsin	286	9	2	0.2814	0.3532	0.3661
breast-cancer-wisconsin	699	10	2	0.1040	0.1216	0.1418

VI. CONCLUSION

In many applications, black-box prediction is not satisfactory, and understanding and handling the data is of critical importance. Typically, approaches useful for understanding of data involve logical rules, evaluate similarity to prototypes, or are based on visualization or graphical methods.

There are several methods proposed for logical rule generation combining different data types (machine learning, fuzzy decision trees, association rules, Bayesian networks, neural networks, pattern recognition). We have selected the more powerful of these algorithms that have been proved from the literature that give better rules and keep the level of interpretability and accuracy in the classification tasks [4], [17], [18].

Induction of FDT is a useful technique to find patterns in data in the presence of imprecision, either because data are fuzzy in nature or because we must improve its semantics. We have proposed the technique to induction of new type of fuzzy decision tree – Stable FDT, which is simple to understand and apply. The use of cumulative information estimations allows precisely estimating mutual influence of attributes. These evaluations are used to analyze group of training instances.

REFERENCES

[1] V. Podgorelec, P. Kokol, B. Stiglic, I. Rozman, “Decision trees: an overview and their use in medicine,” Journal of Medical Systems, vol. 26, pp. 445-463, Oct. 2002.

[2] E.I. Papageorgiou, N.I. Papandrianos, D. Apostolopoulos1, P. Vassilakos, “Complementary use of Fuzzy Decision Trees and Augmented Fuzzy Cognitive Maps for Decision Making in Medical Informatics,” in Proc. 2008 Int. Conf. on BioMedical Engineering and Informatics, Sanya, Hainan, China, pp.888-892.

[3] C. Marsala, “A Fuzzy Decision Tree Based Approach to Characterize Medical Data,” in Proc. 2009 IEEE Int. Conf. on Fuzzy Systems, Jeju Island, Korea, pp.1332-1337.

[4] L. Rokach, O. Maimon, Data Mining with Decision Trees, World Scientific, 2007.

[5] M. U. Khan, J. P. Choi, H. Shin, M. Kim, “Predicting Breast Cancer Survivability Using Fuzzy Decision Trees for Personalized Healthcare”, in Proc. 30th Annual Int. IEEE Conf. on Engineering in Medicine and Biology Society, Vancouver, Canada, 2008, p.5148-5151

[6] V. Levashenko, E. Zaitseva, S. Puuronen, “Fuzzy Classified Based on Fuzzy Decision Tree”, in Proc. 2007 IEEE Int. Conf. on Computer as a tool (EUROCON 2007), Warsaw, Poland, pp.823 – 827.

[7] V. Levashenko, E. Zaitseva, “Usage of New Information Estimations for Induction of Fuzzy Decision Trees”, in H. Yin et al. (Eds.): IDEAL 2002, Lecture Notes in Computer Science LCNS2412, Springer-Verlag, pp. 493-499, 2002.

[8] S. N. Ghazavi, T. W. Liao, “Medical data mining by fuzzy modeling with selected features,” Artificial Intelligence in Medicine, vol. 43, pp. 195–206, 2008.

[9] Information Technology Solution for Healthcare, K. Zielinski M. Duplaga, D. Ingram Ed. London: Springer, 2006.

[10] J.R. Quinlan, “Simplifying decision trees,“ International Journal of Manmachine Studies, N. 27, pp. 221-234, 1987.

[11] H.-M. Lee, C.M. Chen, J.M. Chen, Y.L. Jou, “An Efficient Fuzzy Classifier with Feature Selection Based on Fuzzy Entropy,” Journal of IEEE Trans. on Systems, Man and Cybernetics, Part B:Cybernetics, vol.31(3), pp. 426-432, 2001.

[12] L.Zadeh, “Fuzzy Sets as a basis for theory of possibility,” Fuzzy Sets and Systems, vol.1, pp.3-28, 1978.

[13] Kaufmann A., Gupta. M., Induction to fuzzy arithmetic: theory and applications, New York: Van Nostrand Reinold Co., 1985.

[14] Y. Yuan, M.J. Shaw, “Induction of Fuzzy Decision Trees,” Fuzzy Sets and Systems, vol. 69, pp.125-139, 1995.

[15] http://www.cancer.gov/cancertopics/types/breast

[16] C. Olaru, L. Whenkel, “A Complete Fuzzy Decision Tree Technique,” Fuzzy Sets and Systems, pp.221-254, 2003,

[17] C.Z.Janikow, “Fuzzy Decision Trees: Issues and Methods,” Journal of IEEE Trans. on Systems, Man, and Cybernetics — Part B: Cybernetics, vol. 28, pp.1-14, Jan. 1998.

[18] S.Mitra, K.M.Konwar, and S.K.Pal, “Fuzzy Decision Tree, Linguistic Rules and Fuzzy Knowledge-Based Network: Generation and Evaluation,” Journal of IEEE Trans. on Syst., Man Cybernetics — Part C: Applications and Reviews, vol. 32, pp.328-339, Apr. 2002.

[19] K. Crockett, Z. Bandar, J. O’Shea, “Fuzzification of Discrete Attributes From Financial Data in Fuzzy Classification Trees,” in Proc. 2009 IEEE Int. Conf. on Fuzzy Systems, Jeju Island, Korea, pp.1320-1325.

[20] S.Hettich, C.L.Blake, and C.J.Merz, “UCI Repository of ML Databases. Irvine, CA, University of California”, Dept. of Information and Computer Science, 1998.

[21] O. L. Mangasarian and W. H. Wolberg, “Cancer diagnosis via linear programming”, SIAM News, vol.23, pp 1-18, Sept. 1990.