MODELING FUZZY DECISION TREES FROM FUZZY SETS*

Levashenko Vitaly [1, 3], Puuronen Seppo [2] and Kovalik Stefan [3]

[1] *Dept. of Information Technologies, Belarus State Economic University, Minsk, Belarus*

[2] *Dept. of Computer Science and Information systems, University of Jyväskylä, Finland*

[3] *Dept. of Informatics, University of Žilina, Slovakia*

Introduction

Current electronic data repositories contain enormous amounts of data. These data include also currently unknown and potentially interesting patterns and relations that can be uncovered using knowledge discovery and data mining methods [1]. Commonly supervised machine learning is used, in which there exists a set of training instances represented by a vector of the values of features and a class label.

In this paper we are present our approach that can deal with fuzzy defined data. These data are more accuracy to reflect the real around world. We use a technique to compute *cumulative* information estimations [2]. The use of such estimations allows inducing minimum cost fuzzy decision trees (FDT) based on different criteria of optimality. As example we will introduce three types in this paper: *unordered*, *ordered* and *reliability* FDT.

1. Fuzzy Decision Trees

Quinlan proposed an ID3 algorithm to design crisp decision trees [3]. The first idea of FDT was introduced by Chang and Pavlidis [4]. They presented FDT structure and a search method and described a difference between crisp and FDT. The generalizing ID3 algorithm for fuzzy sets was resulted into Fuzzy ID3 algorithm and its known variants that can be found in [2, 4-11]. The classification task in [5] has nominal or numerical values as input and output attributes. No fuzziness was involved with nominal data. Kosko's fuzzy entropy was used to measure the fuzziness of classification by the neuron [6]. Yuan and Shaw proposed construction a FDT in the process of reducing classification *ambiguity* with accumulated fuzzy *evidences* [7]. This algorithm was developed in [8].

* This investigation has been partially supported by the Belarus Republican Fund of Fundamental Researches and Ministry of Education (Republic of Belarus) and grand of Visiting Researcher of Finland Academy.

2. Fuzzy Decision Trees Induction

A typical classification problem can be described as follows [7-8]. A universe of objects U={u} is described by N training examples and n input attributes **A**={A_1,...,A_n}. Each attribute A_i ($1 \le i \le n$) measures some important feature and is 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 $Cost_i$ is an integrated measure that accounts financial and temporal costs that are required to define the value of the A_i for a certain subject. We will suggest that each object u in the universe is classified by a set of classes $\{B_1,\ldots,B_{m_b}\}$. This set describes by output attribute B.

In this paper we introduce an approach for a sequence of expanded attributes testing, i.e. determination of input attributes' values $\{A_{1i},\ldots,A_{ni}\}$ of a new subject, that allows to accomplish the correct diagnostics. It is obvious that the problem is that the quasioptimal sequence should guarantee as correct diagnostics with a priori defined level of accuracy as minimum cost for accomplishing the tests and procedures.

Example 1. (adopted from [7-8] with some modifications) An object is presented with four attributes: **A**={A_1,A_2,A_3,A_4} and one output attribute B. Each attribute has 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 $B = \{B_1, B_2, B_3\}$.

The membership and cardinality measure of values of these attributes is presented in Table.

We propose new interpretation of Fuzzy ID3, which is based on *cumulative* information estimates [2] for data presented in Table. We realize three types of FDT with different properties: (a) *unordered FDT* (see Fig.1); (*b*) *ordered FDT*, in which to every node of one level associate similar expanded attribute (see Fig.2); (c) *reliability FDT* which more stable for different noise (see Fig.3).

Table. A small 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
$Cost_i$	0,2			0,15			0,45		0,2				
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
$M(A_{ij})$	**6.6**	**5.8**	**3.6**	**6.1**	**5.6**	**4.3**	**5.6**	**10.4**	**6.5**	**9.5**	**4.4**	**4.4**	**7.2**

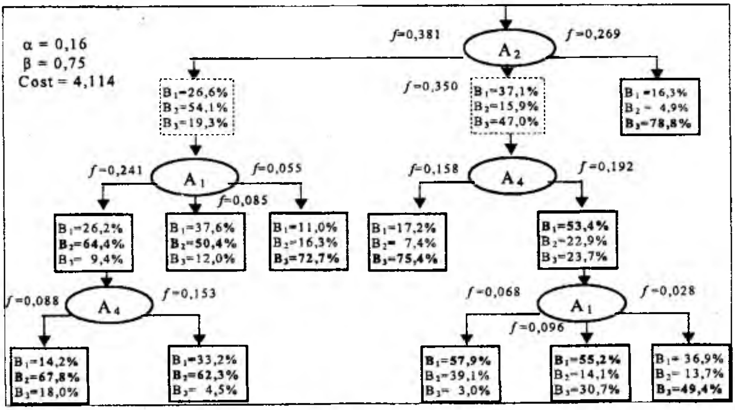

Figure 1. Unordered FDT

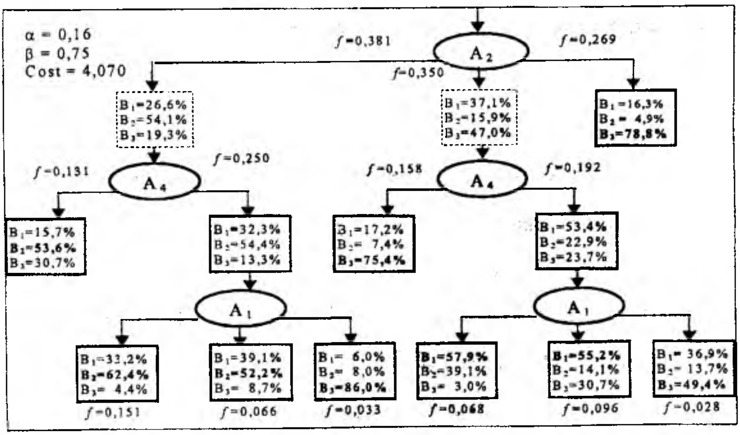

Figure 2. Ordered FDT

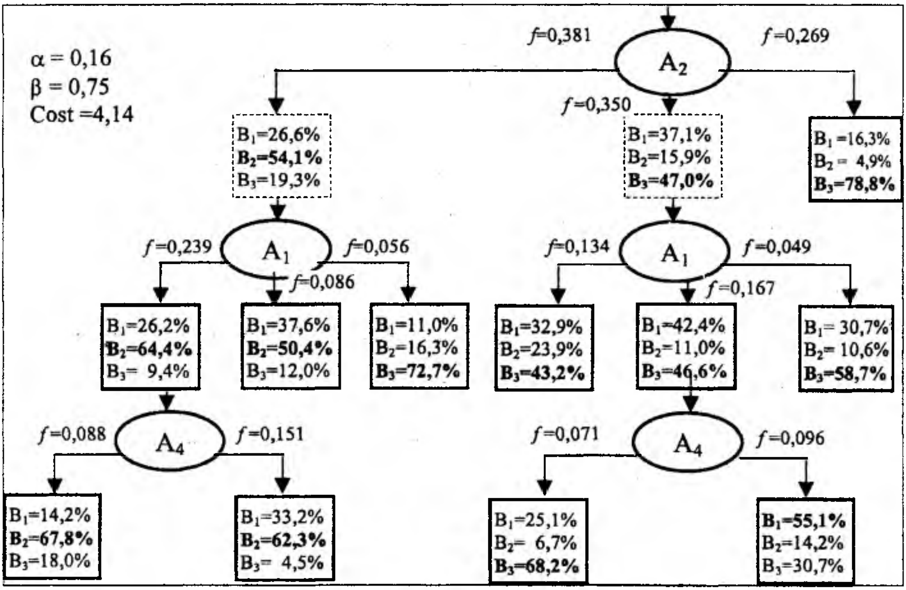

Figure 3. Reliability FDT

4. Conclusion

In this paper we have shown an application of our information estimations [2] in a modeling "greedy" Fuzzy ID3 algorithm for induction of FDT. The use of such estimates allows inducing minimum cost FDT based on different criteria of optimality. We calculated the cost of FDT into considered algorithms.

References

[1] R.M. Goodman, J.W. Miller and P. Smyth, An Information Theoretic Approach to Rule-based Connectionist Expert Systems, *Advances in Neural Information Processing Systems I, Morgan Kaufmann,* pp. 256-263, 1988

[2] V.Levashenko, E.Zaitseva, Usage of New Information Estimations for Induction of Fuzzy Decision Trees. *Proc of the* 3rd *IEEE Int. Conf. on Intelligent Data Engineering and Automated Learning*, UK, 2002

[3] J.R.Quinlan, Induction of Decision Trees, *Machine Learning*, **1**, pp. 81-106, 1986

[4] R.L.P.Chang and T. Pavliddis, Fuzzy Decision Tree Algorithms, *IEEE Transactions on Systems, Man and Cybernetics*, **7**, pp. 28-35, 1977

[5] R.Weber, Fuzzy-ID3: a Class of Methods for Automatic Knowledge Acquisition, *Proc. of the 2nd Int.Conf. on Fuzzy Logic and Neural Networks,* Izuka, Japan, pp. 265-268, 1992.

[6] B.Kosko, Neural Networks and Fuzzy Systems. *Prentice-Hall, Englewood Cliffs*, NJ, 1992

[7] Y.Yuan, M.J. Shaw, Induction of Fuzzy Decision Trees, *Fuzzy Sets and Systems*, **69**, 1995

[8] X.Wang, B.Chen, G.Qian and F.Ye, On the Optimization of Fuzzy Decision Trees, *Fuzzy Sets and Systems*, **112**, pp.117-125, 2000

[9] C.Z.Janikow Fuzzy Decision Trees: Issues and Methods, *IEEE Trans. on Systems, Man, and Cybernetics — Part B*: *Cybernetics*, Vol. 28, **1**, pp.1-14, 1998

[10] S.Mitra, K.M.Konwar, and S.K.Pal, Fuzzy Decision Tree, Linguistic Rules and Fuzzy Knowledge-Based Network: Generation and Evaluation, *IEEE Trans. on Syst., Man Cybernet. — Part C: Applications And Reviews*, Vol. 32, **4**, pp.328-339, 2002

[11] D.S.Yeung, E.C.Tsang, etc, Fuzzy Rule Wining by Fuzzy Decision Tree Induction Based on Fuzzy Feature Subset. *Proc. of the IEEE Int. Conf. on Syst., Man Cybernet*, **4**, pp.599-604, 2002