

LEARNING FUZZY DECISION TREES FROM SEQUENTIAL AND INCOMPLETE DATA

G. Chen, Ruqian Lu, Z. Jin

*Institute of Mathematics, Chinese Academy of Sciences
No. A1, Nansijie, Zhongguancun, Beijing 100080, P.R. China
cg@amss.ac.cn, rqlu@math08.ac.cn, zhijin@math03.math.ac.cn
Tel: +86-10-62554389*

Abstract: In this paper we present a way of constructing and using fuzzy decision trees for machine learning from sequential and incomplete data. We develop a theory and technique for processing such data. At first, we introduce a decision tree with fuzzy attributes and class the properties of fuzzy attributes and classes, where we will introduce the fuzzy division of objects in classes in a different way than that used in the literature. Then we will also introduce properties of fuzzy attribute value to enrich the fuzziness of our decision trees.

Keywords: fuzzy decision trees, tendency detection, sequential and incomplete data

1. INTRODUCTION

Fuzzy decision trees have been used in machine learning and data analysis to solve the problem of incomplete data and noisy data [2, 3, 5]. When using the technique of fuzzy decision trees, the fuzziness of decision-making may be introduced at different points such as: a) The values of attributes may be fuzzy, b) The matching of attribute values may be fuzzy, c) The division of objects in classes may be fuzzy, d) The branching of a node in different arcs may be fuzzy. In literature, the properties b) and d) are often used to define a fuzzy decision tree. For example, Chang and Pavlidis[1] introduced a fuzzy decision tree such that each non-leaf node I has a k -tuple decision function f_i from the sample set X to $[0,1]^k$ and k ordered sons. In

The original version of this chapter was revised: The copyright line was incorrect. This has been corrected. The Erratum to this chapter is available at DOI: [10.1007/978-0-387-35602-0_35](https://doi.org/10.1007/978-0-387-35602-0_35)

M. A. Musen et al. (eds.), *Intelligent Information Processing*

© IFIP International Federation for Information Processing 2002

Janikow's fuzzy decision trees[2], an example can match more than one condition and the example may eventually fall into many of the leaves.

2. FUZZY DECISION TREE

A fuzzy decision tree with the goal classes (g_1, g_2, \dots, g_L) , the attributes (a_1, a_2, \dots, a_n) , and the dictionary of fuzziness $DOF=(C, D, F, M, Q)$, is a tree, called *FDT*, with the following properties: 1) It is finite, 2) Each leaf of the tree corresponds to a membership vector $m=(m_1, m_2, \dots, m_L)$, where L is the number of classes, and where each m_i is a value from $[0,1]$, denoting the degree to which this leaf node belongs to the i^{th} class, 3) Each branch is attached with a fuzzy expression c_i , where c_i belongs to F of *DOF*.

Let there be in total L different classes which form a set G . The membership vector of an object with regard to these classes is a L -dimensional vector, of which each component is a real number in $[0,1]$. Let $ob[i]$ be the i^{th} component of the membership vector of the object ob . This $ob[i]$ denotes the membership degree of ob with regard to the i^{th} class g_i . Further let the appearance maximum $apmax(g_i, X)$ and the appearance minimum $apmin(g_i, X)$ of the i^{th} class g_i with regard to the object set X be

$$apmax(g_i, X) = \text{Max}\{ob[i] \mid ob \text{ is an object of } X\}$$

$$apmin(g_i, X) = \text{Min}\{ob[i] \mid ob \text{ is an object of } X\}$$

then the following sum:

$$H(X,G) = \sum\{apmax(g_i, X) \times \log apmin(g_i, X) \mid \text{for all } i\}$$

is called the fuzzy entropy of the set X with regard to the set G of classes.

3. TENDENCY DETECTION WITH FDT

The length of a finite sequence a , $\text{length}(a)$, is the number of its elements minus one. The average height of the sequence a , $\text{high}(a)$, whose elements are real numbers, is the mean value of these values.

Let $SVL=(a_0, a_1, \dots, a_N)$ be a sequence. We say that a_1 is its head and a_N its tail. Further we define the tendency of change *TOC*, of the sequence *SVL* as follows:

$$TOC = (a_N - a_0) / N$$

We call the tendency positive if *TOC* is positive, otherwise we call it negative. Length, average height and tendency of change are the three parameters which characterize the state of a sequence. We use a unified name tendency for all of the three together.

Now we have transformed the problem of classifying exact value sequences into one of classifying fuzzy value sequences. The question remaining is how to compare vectors (sequences) of unequal length and how to use them in the decision procedure.

4. CONSTRUCTING FUZZY DECISION TREES

Since one-way adjusted value sequences are rather frequently met in the practice. We are going to first process this kind of incomplete data. Here we also allow fuzzy values. Since the fuzzy value sequences are usually not of equal length, we will not use the method of vector comparison or vector clustering. Rather, we will try to construct a new kind of fuzzy decision trees which also apply to one-way-adjusted fuzzy value sequences. However, before we go into the details, we should differentiate between two issue: (1) To construct a fuzzy decision tree by using incomplete data. (2) To test an example consisting of incomplete data. In the following we will only discuss the issue of constructing a decision tree for one-way adjusted data.

Algorithm 1. (Construction of Fuzzy Decision Trees for Pre-adjusted Sequences).

1. Let X be the set of all examples and $ihr > 0$ be some threshold value. Each example is a fuzzed data sequence as described in algorithm 1. Let A and B two empty sets of open and closed nodes, respectively.
2. Construct a decision tree T with only one node, i.e. the root (X, Z) where X is the set of all examples, Z is the length of the longest fuzzy value sequence among all examples, i.e., $Z = \max \{ |seq| \mid seq \text{ is the fuzzy value sequence of some example from } X \}$
Put (X, Z) in the set A .
3. Let (X', Z') be an open leaf node. If all examples of X' belong to the same set of classes (for a precise meaning of this statement see step 3 of algorithm 1), or if Z equals to zero, then fetch this node from A and put it into B . Call the vector

$$M = \text{Mean}\{m_i \mid m_i \text{ is the membership vector of some example from } X\}$$
as the membership vector of (X, Z) . The average is taken for each component separately. Do the same for all nodes of such kind.
4. Stop the algorithm if A becomes empty.
5. Fetch a node (X', Z') from A and put it into B . Let $i = Z - Z' + 1$.
6. If there is no non-empty subset X'' of X' , where no sequence has an i -th item, then go to step 9. Otherwise build the vector

$$M = \text{mean}\{m_i \mid m_i \text{ is the membership vector of some example from } X\}$$
The average is separately taken for each component.

7. If $X'=X''$ then define M be its membership vector and Go to step 3.
8. Otherwise, call M the complement vector of (X', Z') . Let $X'=X'-X''$.
9. Let the different fuzzy values *FR* of the *i*-th items of all examples of the set X' be (d_1, d_2, \dots, d_G) . Construct *G* branches from the node (X', Z') , marked with the label (d_1, d_2, \dots, d_G) , respectively. Call the endpoints of these branches (X^j, Z^j-1) respectively, for $j=1, \dots, G$, where the *i*th item of the fuzzy value sequence of each example from X^j is equal to d_j . We put all of them in *A*.
10. Go to step 3.

5. CONCLUSION

We have present here our work on how detect temporal tendencies in form of fuzzy decision trees. A major part of this research has been contributed to the solution of the problem, that the temporal knowledge and data we collect is often incomplete.

6. REFERENCE

- [1] Chang, R.L.P. and Pavlids, T. Fuzzy decision trees. In Proceedings of the IEEE Transactions on Systems, Man, and Cybernetics, pages 360-367, 1976.
- [2] Janikow, C.Z.. Fuzzy processing in decision trees. In Proceedings of the sixth International Symposium on Artificial Intelligent, pages 360-367, 1993.
- [3] Janikow, C.Z.. Fuzzy decision trees, issues and methods. IEEE Transactions on Systems, Man, and Cybernetics, 1997.
- [4] Lu,R. and Jin,Z. Domain modeling-based software engineering, Kluwer Academic Publishers. August 2000
- [5] M.Umato and et al. Fuzzy decision trees by fuzzy ID3 algorithm and its application to diagnosis systems. In Proceedings of the third IEEE Conference onFuzzy Systems, pages 2113-2118, 1994.