Information Sciences 465 (2018) 425–438

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

# Tolerance rough fuzzy decision tree

Junhai Zhai[a,*], Xizhao Wang[b], Sufang Zhang[c], Shaoxing Hou[d]

[a] Key Lab. of Machine Learning and Computational Intelligence, College of Mathematics and Information Science, Hebei University, Baoding, Hebei 071002, China
[b] College of Computer Science and Software, Shenzhen University, Shenzhen 518060, China
[c] Hebei Branch of Meteorological Cadres Training Institute, China Meteorological Administration, Baoding, China
[d] College of Computer Science and Technology, Hebei University, Baoding, Hebei 071002, China

ARTICLE INFO

*Article history:*
Received 8 September 2016
Revised 19 April 2018
Accepted 5 July 2018

*Keywords:*
Fuzzy decision tree
Rough set
Tolerance rough set
Rough fuzzy set
Tolerance rough fuzzy set

ABSTRACT

Fuzzy decision tree (FDT) is an extension of decision tree. Fuzzy classification rules can be extracted by FDT from fuzzy decision tables with fuzzy conditional attributes and fuzzy decision attributes. However, it is very time consuming for fuzzifying conditional attributes, and fuzzification of conditional attributes will inevitably lead to information loss. In order to deal with this problem, based on tolerance rough fuzzy set, this paper proposed an algorithm named TRFDT (Tolerance Rough Fuzzy Decision Tree) and theoretically proved that the proposed algorithm is convergent with a very large probability. TRFDT can directly handle fuzzy decision tables with continuous-valued conditional attributes and fuzzy decision attributes. Accordingly, TRFDT has fast learning speed and good generalization ability, which have been experimentally proved by comparing TRFDT with two state-of-the-art approaches fuzzy ID3 and FDT-YS.



## 1. Introduction

Fuzzy decision tree (FDT) [29] is an extension of decision tree [24]. Fuzzy classification rules can be extracted by FDT from fuzzy decision tables with fuzzy conditional attributes and fuzzy decision attributes. Because fuzzy classification systems based on fuzzy decision tree are generally robust, FDT have been widely and successfully applied to many fields [25], such as decision-making [9,50,51], classification and prediction [38,45,49], biological information processing [15,16], etc. Fuzzy ID3 [29] and the fuzzy decision tree algorithm proposed by Yuan and Shaw [39] (denoted by FDT-YS for convenience in this paper) are two famous fuzzy decision tree algorithms. Fuzzy ID3 was directly extended from ID3 [24] by replacing information entropy with average fuzzy classification entropy to select expanded attributes. Different from fuzzy ID3, FDT-YS [39] employed ambiguity as heuristic to select expanded attributes, the termination condition for leaf node is same as fuzzy ID3. Along with this technology route, many fuzzy decision tree algorithms were proposed by different researchers. Based on fuzzy rough set, Zhai [43] proposed a fuzzy rough decision tree algorithm which cleverly combined the roughness of knowledge and fuzziness of data. Compared with fuzzy ID3, the testing accuracy can be improved by this algorithm. Olaru and Wehenkel [21] proposed a novel fuzzy decision tree named soft decision tree which combined tree-growing and pruning, to determine the structure of the soft decision tree, refitting and backfitting techniques were used to improve the generalization capability of soft decision tree. Jang [12] extended CART algorithm to fuzzy environment, and proposed

* Corresponding author.
*E-mail address:* mczjh@hbu.cn (J. Zhai).

https://doi.org/10.1016/j.ins.2018.07.006

fuzzy CART. The fuzzy CART can produce more comprehensible fuzzy rules. Lertworaprachaya et al. [17] extended the fuzzy ID3 algorithm to interval-valued conditional attributes. They represented fuzzy membership values as intervals to model uncertainty and employed the look-ahead based fuzzy decision tree induction method to construct decision tree. They also investigated the significance of different neighbourhood values and defined a new parameter insensitive to specific data sets using fuzzy sets. Zaitseva and Levashenko [41] proposed a FDT based method for construction of structure function, which was used to represent the correlation of the system performance level and the states of its components. Based on Hadoop MapReduce, Segatori et al. [26] proposed a distributed learning scheme for generating fuzzy decision trees, the proposed scheme is suitable for managing big data sets even with a modest commodity hardware support.

Except these algorithms, there are also some other algorithms proposed from different technology routes. For instance, Based on axiomatic fuzzy sets, Liu et al. [19] proposed a fuzzy decision tree algorithm for extracting fuzzy classification rules, which can be applied to data sets with mixed data type attributes. Zeinalkhani and Eftekhari [42] proposed a two steps algorithm for constructing fuzzy decision tree. In the first step, discretization divides domain of continuous attributes to several partitions, and then, in the second step, an fuzzy membership degree is defined on each partition. Finally a fuzzy decision tree is generated. Wang et al. [34] proposed fuzzy rule based decision trees. In contrast with traditional axis-parallel decision trees in which only a single feature is taken into account at each node, each node of the proposed decision trees involves multiple features. Pedrycz and Sosnowski [23] proposed a clustering fuzzy decision tree algorithm, which introduced the idea of granular computing into the process of induction of fuzzy decision tree. Based on case-based reasoning, Chang et al. [2] proposed a fuzzy decision tree algorithm and they applied the proposed algorithm to data classification. Wang et al. [32] proposed an algorithm for generating fuzzy decision trees with carefully selected samples, the proposed algorithm can significantly improve the generalization ability of fuzzy decision trees. Wang et al. [31] conducted a further investigation on the relationship between the uncertainty and generalization ability of fuzzy learning algorithms, and obtained very valuable conclusion: the classifier with higher uncertainty outputs has better performance for complex boundary problems. Recently, Wang [35] edited a special issue on learning with uncertainty. In addition, Wang et al. also studied the applications of uncertainty in machine learning. For instance, fuzziness based semi-supervised learning and its application in intrusion detection was studied in [1]. Fuzziness based nonlinear regression was studied in [10]. OWA operator based link prediction ensemble and its application in social network data mining was investigated in [8]. Fuzzy rough set based feature selection methods were investigated in [33,36,48]. A comparative study on heuristic algorithms for generating fuzzy decision trees was given in [30], an excellent survey of fuzzy decision tree algorithms can be found in [13].

These algorithms mentioned above can only induce fuzzy decision trees from fuzzy decision tables with fuzzy-valued conditional attributes and fuzzy-valued decision attribute. When these algorithms are applied to fuzzy decision tables with continuous-valued conditional attributes and fuzzy-valued decision attribute, it is inevitable for these algorithms to fuzzify the continuous-valued conditional attributes, but it is difficult to determine the fuzzy membership degree. Based on our previous work [46], in this paper, an induction algorithm of fuzzy decision tree named TRFDT is proposed, TRFDT can directly induce fuzzy decision tree from fuzzy decision tables with continuous-valued conditional attributes and fuzzy-valued decision attribute. The degree of tolerance rough fuzzy dependency [46,47] is employed to select expanded attributes, the Luca-Termini fuzzy entropy [20] is employed to select the optimal cut, and the Kosko fuzzy entropy [14] is used as termination condition for leaf nodes. We theoretically proved that the proposed algorithm is convergent with a very large probability. The experimental results and analysis including statistical analysis verified that the proposed algorithm TRFDT is effective and efficient.

The remainder of this paper is structured as follows. Section 2 presents the preliminaries, including rough set, tolerance rough set, rough fuzzy set. Tolerance rough fuzzy set and tolerance rough fuzzy decision tree are presented in Section 3. The theoretical proof of the convergence of the proposed algorithm is also presented in this section. An example is provided in Section 4 to illustrate the generation of tolerance rough fuzzy decision trees by the proposed algorithm. Experimental results and statistical analysis are given in Sections 5 and6 concludes this paper.

## 2. Preliminaries

In this section, we briefly review preliminaries including rough set [22], tolerance rough set [27], rough fuzzy set [4]. Some new extended models of rough set can be found in [5–7,28,37].

### 2.1. Rough set

In this paper, we discuss problems in the framework of classification. Let $DT = (U, A \cup C)$ be a decision table with symbolic-valued conditional attributes. $U = \{x_1, x_2, \ldots, x_n\}$, $A = \{a_1, a_2, \ldots, a_d\}$, the instances in $U$ are categorized into $k$ classes: $C_1, C_2, \ldots, C_k$, i.e., $U/C = \{C_1, C_2, \ldots C_k\}$. Let $x \in U$ and $R$ is an equivalence relation induced by a subset of $A$, the equivalence class containing $x$ is given by:

$$[x]_R = \{y | xRy\}. \tag{1}$$

Given a decision table, for arbitrary target concept $C_i \in U/C (1 \le i \le k)$, the lower approximation and the upper approximation of $C_i$ with respect to $R$ are defined by

$$\underline{R}(C_i) = \{[x]_R | [x]_R \subseteq C_i\}. \tag{2}$$