Construct Fuzzy Decision Trees Based on Roughness Measures

Mohamed A. Elashiri[1], Hesham A. Hefny[2], and Ashraf H. Abd Elwhab[3]

[1] Computer Sciences Department, Academy of Specialized Studies, Cairo, Egypt
ashiry@aun.edu.eg
[2] Computer Sciences and Information Department, Institute of Statistics and Research, Cairo University, Egypt
hehefny@ieee.org
[3] Computer Sciences and Systems Department, Electronics Research Institute, Cairo, Egypt
awahab@mcit.gov.eg

Abstract. Data mining is a process of extracting useful patterns and regularities from large bodies of data. Decision trees (DT) is one of data mining techniques used to deal with classical data. Fuzzy Decision Trees (FDT) is generalization of crisp decision trees, which aims to combine symbolic decision trees with approximate reasoning offered by fuzzy representation. Given a fuzzy information system (FIS), fuzzy expanded attributes play a crucial role in fuzzy decision trees. In this paper the problem is slowness and complexities of the fuzzy decision trees, but its rules are more accurate. Our target is to simplify computational procedures and increase the accuracy rules or to keep the high grade of accuracy and to select an efficient criterion to select fuzzy expanded attributes based on rough set theory.

Keywords: Fuzzy set, Rough set, Fuzzy decision tree, Accuracy measure, Roughness measure.

1 Introduction

Decision Tree Induction (DTI), one of the Data Mining classification methods, is used in this research for predictive problem solving in analyzing patient medical track records. In this paper, we have extended the concept of DTI dealing with meaningful fuzzy labels in order to express human knowledge for mining fuzzy association rules with using fuzzy rough technique [1]. Theories of fuzzy sets and rough sets are generalizations of classical set theory for modeling vagueness and uncertainty [2], [3]. Rough sets are the results of approximating crisp sets using equivalence classes, in other words, in traditional rough set approach, the values of attributes are assumed to be nominal data, i.e. symbols. Fuzzy set theory deals with the ill definition of the boundary of a class through a continuous generalization of set characteristic functions. In many applications, however, the attribute values can be linguistic terms (i.e. fuzzy sets), for example, the attribute ‘‘height’’ may be given the values of ‘‘high’’, ‘‘mid’’, and ‘‘low’’, then traditional rough set approach would treat these

V.V. Das and J. Stephen (Eds.): CNC 2012, LNICST 108, pp. 199–207, 2012.

values as symbols, thereby some important information is included in these values such as the partial ordering and membership degrees is ignored, which means that the traditional rough sets approach cannot effectively deal with fuzzy initial data e.g. linguistic terms [4], [5], [6], [7]. The rest of the paper is organized as follows: Section 2 previous research efforts in fuzzy decision trees. Section 3 proposed an algorithm of generating fuzzy rough decision trees. Section 4 gives our experimental results. Section 5 concludes. Section 6 the future work.

2 Previous Research Efforts in Fuzzy Decision Trees

Fuzzy decision trees can process data expressed with symbolic, numerical values and fuzzy terms. The apparent advantage of fuzzy decision trees is that they use the same routines as symbolic decision trees but with fuzzy representation [8]. This allows for utilization of the same comprehensible tree structure for knowledge understanding and verification. Rough set can reduce the fuzzy attribute and lead to simplify of computational rather than computational of fuzzy set, fuzzification of rough sets has two quantifies are [9]: uncertainties simultaneously or vagueness through the fuzzification of real valued attributes and ambiguity through rough approximations [10]. Next subsections displays summary about some heuristics as related work using fuzzy and rough concepts to generate FDT.

2.1 FDT Bases on Jaccard Similarities Measure

Using Jaccard similarities measure to select expanded attribute and combine between each sub-attributes of all attributes with each sub-classes. Moreover aggregation operators such as *MIN*, *MAX*, *OWA*, *WA* must be used to aggregate from depth 0 to depth 1 also, previous combination will be repeating until complete the decision tree. Unattractive reasons are; more than one tree generating in each level to aggregate with others in next depths that lead to construct many of trees. Also, the algorithms of pruning are needed to eliminate some of generated trees [6].

2.2 FDT Bases on Rough Attribute Reducts

According to Fuzzy Iterative Dichotomiser-3 algorithm (FID3) and depend on rough attribute reducts at each time more than one of fuzzy decision trees will be constructed. Many of fuzzy decision trees according to each set of fuzzy reduct attribute will be created also, lot of rules from each tree may be deduced and others algorithms will be needed [7], [8], [11].

2.3 FDT Bases on Dependency Degree Measure

Constructing fuzzy decision trees using attribute selection criteria based on dependency degree of rough set theory, dependency degree of partition. But in the next levels hasn't any constrains to stop the growth of tree, so if dependency degree value of next partition is greater than dependency degree value of previous partition that implies to large size of tree [12].

3 Proposed Algorithm of Generating Fuzzy Rough Decision Tree

The main objective of this paper is to enhance the construction of fuzzy decision trees by applying the integration between theories of fuzzy set and rough set, which has been applied successfully in too many fields such as machine learning, pattern recognition and image processing. This integration lead to efficient rules generation and smaller decision trees which deal with fuzzy or crisp data sets, try to improve the accuracy and also, to overcome the drawbacks of both fuzzy decision trees and rough crisp decision trees. The drawbacks of rough decision trees are, it deals only with data in classical or crisp forms and it cannot effectively deal with fuzzy initial data e.g. linguistic terms. On the other hand, the drawbacks of fuzzy decision trees are: from a computational point of view, with increased size of tree comes increased the computational complexity, and from point of view about pruning which decrease the size of large trees. Using rough set theory is a good way to determine relation within dataset sample that may lead to determine the core of data by its measures such as roughness degree [16].

Consider a non leaf node S consisting of n attributes $S=\{F^1,...,F^n\}$ to be selected for each $1\le k\le n$, the attribute F^k takes m_k values of fuzzy subsets $F^k=\{F_1^{\ k},...,F_{mk}^{\ k}\}$, the fuzzy classification is $F^C=\{F^C_{\ 1},..,F^C_{\ mc}\}$ and universe of discourse is $U=\{X_1,..,X_N\}$, as shown in table 1.

Table 1. Structure of fuzzy information System

U	F^1			.	F^k			.	F^{Class}		
	$F^1_{\ 1}$	....	$F^1_{\ m1}$	.	$F^k_{\ 1}$	...	$F^k_{\ mk}$	.	$F^C_{\ 1}$	....	$F^C_{\ mc}$
X_1	a_{11}^1	...	$a_{1m_1}^1$		a_{k1}^1	...	$a_{km_k}^1$		a_{C1}^1		$a_{Cm_c}^1$
:	:	:	:		:	:	:		:	:	:
X_N	a_{11}^N	...	$a_{1m_1}^N$		a_{k1}^N	...	$a_{km_k}^N$		a_{C1}^N		$a_{Cm_c}^N$

Our paper introduces modify of roughness measure formula, where the roughness measure was defined only on rough set theory. Next section presents a new formulation of roughness measure in fuzzy rough set theory and roughness measure of fuzzy partition.

3.1 New Criterion for Selecting Expanded Attribute

Pawlak proposed two numerical measures for evaluating uncertainty of a set are accuracy and roughness [13], definitions of accuracy measure and roughness measure was given, and suggested form in fuzzy rough set which is as follows [14]:

Definition 1. The accuracy measure is equal to completeness of knowledge about the given object set *x* is defined by the ratio of cardinality of lower and upper approximation set of x as [16], [17], [18], [19]:

$$\alpha_R(x) = \frac{|\underline{R}(x)|}{|\overline{R}(x)|}. \tag{1}$$

Definition 2. The roughness measure represents the degree of incompleteness of knowledge about the given object set *x* is defined as [16]:

$$\psi_R(x) = 1 - \alpha_R(x). \tag{2}$$

Definition 3. Fuzzy rough set is a tuple from lower approximation membership degree and upper approximation membership degree of class $\underline{l}$ through F^k_j, both are define in Eq. (3) and Eq. (4) as [26]:

$$\mu_{\underline{l}}\left(F_j^k\right) = \underset{\forall i \in u}{Inf} \max\left\{1-\mu\left(F_j^k(x^i)\right), \mu_l(y^i)\right\}. \tag{3}$$

$$\mu_{\overline{l}}\left(F_j^k\right) = \underset{\forall i \in u}{Sup} \min\left\{\mu\left(F_j^k(x^i)\right), \mu_l(y^i)\right\}. \tag{4}$$

For arbitrary class $\underline{l}$ from fuzzy classification $F^C=\{F^C_1,F^C_2,\ldots,F^C_{mc}\}$ and fuzzy partitions F^k_j for $1\leq j\leq m_k$ and $1\leq k\leq n$.

Definition 4. The accuracy measure is equal to completeness of knowledge about the fuzzy partition F^k_{j} is defined by the ratio of lower approximation membership degree and upper approximation membership degree of class $\underline{l}$, suggested forms of the accuracy measure in fuzzy data set and fuzzy rough set is [20]:

$$\alpha_l\left(F_j^k\right) = \frac{\mu_{\underline{l}}\left(F_j^k\right)}{\mu_{\overline{l}}\left(F_j^k\right)}. \tag{5}$$

Definition 5. The roughness measure in the fuzzy form equal to incompleteness of knowledge about the fuzzy partition F^k_j, and suggested form of the roughness measure in fuzzy form is:

$$\psi_l\left(F_j^k\right) = 1 - \alpha_l\left(F_j^k\right). \tag{6}$$

3.2 Special Issues to Construct Fuzzy Decision Trees

This section presents special issues to construct fuzzy decision trees such as attribute selection criteria, stopping criteria and splitting of tree.

3.2.1 Select Root of the Tree

Compute roughness measure to each sub of attributes with all subclass of class attribute using Eq. (6), then compute roughness measure with each attribute as:

$$\psi(F^k) = \sum_{j=1}^{m_k} \left(\frac{M\left(F_j^k\right)}{\sum_{j=1}^{m_k} M\left(F_j^k\right)} \right) \psi\left(F_j^k\right). \tag{7}$$

The term $M(F^k_j) / \sum^{mk}_{j=1} M(F^k_j)$, is the weights which represent the relative size of subset F^k_j in F. The attribute with minimum value of roughness measure $\psi(F)$ will be selected as a root of fuzzy decision tree i.e. $Root=Min_k(\psi(F^k)) \in S_{root}$, $1 \le k \le N$.

3.2.2 Stopping Criteria

Degree of truth level and criterion of split are the two criteria we have when stopping growth of tree, this section illustrates the degree of truth level as:

Definition 6. Certainty factor or the degree of truth of the rule "If F^k_i Then F^C_j", measure of ρ (F^k_i, F^C_j) can defined as [15], [20]:

$$\beta_{calc} = \rho\Big(\bigcap_{\substack{i \text{ of } k \text{ in the same} \\ \text{branch}}} F_i^k, F_i^C \Big) = \frac{M\left(\mu\left(F_i^k(u)\right) \cap \mu\left(F_i^c(u)\right)\right)}{M\left(\mu\left(F_i^k(u)\right)\right)}. \tag{8}$$

One of the criteria to stop growth of a branch in the tree is based on user defined threshold $\beta_{defined}$, for leaf selection and each branch will undergo a leaf selection test, which is calculation of β_{Calc}. *If* truth level of classifying into one class is greater then $\beta_{defined}$ *then* terminate the branch as a leaf. Selection of $\beta_{defined}$ is depends on the problem to be solved.

3.2.3 Splitting Criteria

Criteria of splitting depend on roughness measure of fuzzy partition, which defined as following:

Definition 7. Roughness measure of fuzzy partition $\psi(F^v | F^k_j)$ is the weight average of roughness measure of fuzzy partition $F^v \in F^P = S\text{-}S_{root}$ on fuzzy evidence F^k_j is defined as:

$$\psi(F^v | F_j^k) = \sum_{l=1}^{m_v} \left(\frac{M(F_l^v \cap F_j^k)}{\sum_{j=1}^{m_k} M(F_l^v \cap F_j^k)} \right) \psi(F_l^v | F_j^k). \tag{9}$$

Definition 8. Fuzzy lower approximation membership degree of class $\underline{l}$ through $F^v \cap F^k_j$ is defined as:

$$\mu_{\underline{l}}\left(F^v \cap F^k_j\right) = \inf_{\forall i \in u} \max \left\{1 - \mu\left(F^v(x^i) \cap F^k_j(x^i)\right), \mu_l(y^i)\right\}. \tag{10}$$

Definition 9. Fuzzy upper approximation membership degree of class $\bar{l}$ through $F^v \cap F^k_j$ is defined as:

$$\mu_{\bar{l}}\left(F^v \cap F^k_j\right) = \sup_{\forall i \in u} \min \left\{\mu\left(F^v(x^i) \cap F^k_j(x^i)\right), \mu_l(y^i)\right\}. \tag{11}$$

If $\psi(F^v \cap F^k_j) \le \psi(F^k_j)$ for any $F^v \in F^p$, where $F^d = \{F^1, F^2, \ldots, F^L\} = S - S_{root}$, then select the attribute F^v with smallest value of roughness measure is a new decision node from of F^k_j branch otherwise terminate F^k_j branch as a leaf with the highest truth level β_{Calc}.

3.3 An Algorithm to Generate Fuzzy Decision Tree

This section introduces the algorithm for generating fuzzy decision tree; the induction process consists of the following steps: consider a non-leaf node S consisting of n attributes $F^1, \ldots, F^n$, to be selected for each k ($1 \le k \le n$), the attribute F^k takes m_k values of fuzzy subsets, $F_1^k, \ldots\ldots F_m^k$, as shown in table 2.

Table 2. Proposed Algorithm to construct fuzzy decision tree

Input: FIS

Output: Group of fuzzy classification rules extracted from the generated fuzzy decision tree T

Step 1: Measure the roughness associated with each attribute and selects the attribute with the smallest value of roughness measure $\psi_l(F^k_j)$ as the root decision node, which implies to pure classification using Eq. (6).

Step 2: Delete all empty branches of the decision node and for each nonempty branch of the decision node, calculate the truth level of classifying all objects within the branch into each class as a leaf.

If truth level of classifying into one class is above a given threshold β from Eq. (8).

Then terminate the branch as a leaf.

Else, investigate

If an additional attribute will further partition the branch and further reduce the value of roughness measure.

Then select the attribute with smallest value of objective function as a new decision node from the branch from Eq. (9).

Else terminate this branch as a leaf and at the leaf; all objects will be labeled to one class with the highest truth level Eq. (8).

Step 3: Repeat *step* 2 for all newly generated decision nodes until no further.

Step 4: Extracting fuzzy classification rules from the fuzzy decision tree T.

4 Experiments

The effectiveness of our technique is demonstrated through numerical experiments in the environment of C#. Our experiments select 5 datasets are from UCI [21], as shown in table 3. Training data, testing data and numbers of rules are summarizes in tables 4, 5,6,7,8 with different fuzzy decision tree algorithms such as Yuan's method or Fuzzy Decision Trees based on measure of Ambiguity (FDTA) [18], [22], [23], [24] FID3 [10], [15], [18], [24] Fuzzy Decision Trees based on measure of Dependency Degree (FDTDD) [12], [14], [25] and our proposed Fuzzy Decision Trees based on Roughness measure (FDTR).

Table 3. Features of data sets

Datasets	Samples	Attributes	Classes
Quinlan's Weather	16	4	3
Wisconsin Breast Cancer	699	9	2
Monk_1	432	7	2
Iris	150	4	3
Pima Diabetes	768	8	2

Table 4. Dataset of Quinlan's

	FDTA	FID3	FDTDD	FDTR
N. Rules	13	16	11	5
Training Accuracy	69.2	69	85	86.1

Table 5. Dataset of Monk_1 Problem

	FDTA	FID3	FDTDD	FDTR
N. Rules	58	59	38	30
Training Accuracy	95	95	94.6	95
Testing Accuracy	93.02	94	94	94.1

Table 6. Dataset of Pima Indian Diabetes

	FDTA	FID3	FDTDD	FDTR
N. Rules	61	64	53	51
Training Accuracy	88.5	88.48	88.5	88.6
Testing Accuracy	89.2	88.9	88	89.3

Table 7. Dataset of Iris Problem

	FDTA	FID3	FDTDD	FDTR
N. Rules	8	9	6	6
Training Accuracy	91.9	91.2	92	92
Testing Accuracy	95.2	96	96.3	96.3

Table 8. Dataset of Wisconsin Breast Cancer

	FDTA	FID3	FDTDD	FDTR
N. Rules	100	102	99	95.9
Training Accuracy	95.2	95.1	95.3	95.6
Testing Accuracy	95.6	95.6	95	96.2

5 Conclusions

New technique have been presented to construct fuzzy decision tree from both crisp and fuzzy dataset depend concepts from rough set theory, the main proposed is simplify computational procedures this was achieved from using simple set "rough set theory" and increase the accuracy rules or keep of the high grade of accuracy, where process of extract rules from huge dataset using FDTA, FID3 and FDTDD be very complex computational procedures rather than complex computational procedures from our proposed FDTR. Our proposed method takes both fuzziness and roughness existing in an information system into consideration; in addition numerically experimental results statistically verify the effectiveness of our proposed method. The results of our technique are found to be smaller than other famous algorithms in tree size and consequently better generalization (test) performance.

6 Future Work

Try to implement this heuristic on dataset from Geographical Information System (GIS) or GIS maps to extract information which determine the accurate position of gold mines, metal mines and oil drilling without drilling.

References

1. Yao, Y.Y.: A comparative study of Fuzzy Sets and Rough Sets. Information Sciences 109(1-4), 227–242 (1998)
2. Zhi, W.W., Mi, J.S., Zhang, W.X.: Generalized fuzzy Rough Sets. Information Sciences 151, 263–282 (2003)
3. Wang, X.Z., Tsangb, E.C.C., Zhao, S., Chen, D., Yeung, D.S.: Learning Fuzzy Rules from Fuzzy Samples based on Rough Set Technique. Information Sciences 177, 4493–4514 (2007)
4. Wang, T., Li, Z., Yan, Y., Chen, H.: A Survey of Fuzzy Decision Tree Classifier Methodology. In: Fuzzy Information and Engineering (ICFIE). ASC, vol. 40, pp. 959–968 (2007)
5. Ning, Y., Tianruii, L., Jing, S.: Construction of Decision Trees based Entropy and Rough Sets under Tolerance Relation. In: International Conference on Intelligent Systems and Knowledge Engineering, ISKE (2007)
6. Huang, Z., Gedeon, T.D., Nikravesh, M.: Pattern Trees Induction: A New Machine Learning Method. IEEE Transactions on Fuzzy Systems 16(4) (2008)

7. Wang, X.Z., Zhai, J.H., Lu, S.X.: Induction of Multiple Fuzzy Decision Trees based on Rough Set Technique. Information Sciences 178, 3188–3202 (2008)
8. Wang, X.Z., Zhai, J.H., Zhang, S.F.: Fuzzy Decision Tree based on the Important Degree of Fuzzy Attribute. In: Proceedings of the Seventh International Conference on Machine Learning and Cybernetics, Kunming, pp. 511–516 (2008)
9. Ruying, S.: Data Mining Based on Fuzzy Rough Set Theory and Its Application in the Glass Identification. Modern Applied Science 3(8) (2009)
10. Janikow, C.Z.: Fuzzy Decision Forest. In: NAFIPS 22nd International Conference of the North American Fuzzy Information Processing Society, pp. 480–483 (2003)
11. Qing, H.H., Ming, Y.G., Da, R.Y.: Construct Rough Decision Forests based on Sequentially Data Reduction. In: Proceedings of the Fifth International Conference on Machine Learning and Cybernetics, pp. 2284–2289 (2006)
12. Rajen, B., Gopal, M.: FRCT: Fuzzy-Rough Classification Trees. Pattern Analysis & Applications 11(1), 73–88 (2008)
13. Pawlak, Z., Skowron, A.: Rudiments of Rough Sets. Information Sciences 177, 3–27 (2007)
14. Zhai, J.: Fuzzy decision tree based on fuzzy-rough technique. Soft Computing 15(6), 1087–1096 (2011)
15. Xue, L., Xiao, H.Z., Dong, D.Z.: Four Matching Operators of Fuzzy Decision Tree Induction. In: Proceedings of the Fourth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2007), vol. 4, pp. 674–678 (2007)
16. Liang, J., Wang, J., Qian, Y.: A New Measure of Uncertainty based on Knowledge Granulation for Rough Sets. Information Sciences 179, 458–470 (2009)
17. Chen, Y.L., Hu, H.W., Tang, K.: Construct a Decision Tree from Data with Hierarchical Class Label. Expert Systems with Applications 36, 4838–4847 (2009)
18. Hefny, H.A., Ahmed, S.G., Abdel Wahab, A.H., Elashiry, M.: Effective Method for Extracting Rules from Fuzzy Decision Trees based on Ambiguity and Classifiability. Universal Journal of Computer Science and Engineering Technology 1(1), 55–63 (2010)
19. Jensen, R., Shen, Q.: Fuzzy-Rough Feature Significance for Fuzzy Decision Trees. In: Proceedings of the 2005 UK Workshop on Computational Intelligence, pp. 89–96 (2005)
20. Elashiry, M.A., Hefny, H.A., Abdel Wahab, A.H.: Fuzzy Rough Decision Tree based on Rough Accuracy Measure and Fuzzy Rough Set. In: Proceedings of the 1st International Conference on Advanced Computing and Communication (ICACC 2010), pp. 134–138. Amal Jyothi College of Engineering, Kerala (2010)
21. Blake, C.L., Merz, C.J.: UCI Repository of machine learning databases, `http://archive.ics.uci.edu/ml/`
22. Yuan, Y., Shaw, M.J.: Induction of fuzzy decision trees. Fuzzy Sets and systems 69, 125–139 (1995)
23. Wang, X.Z., Gao, X.H.: A Research on the Relation Between Training Ambiguity and Generalization Capability. In: International Conference on Machine Learning and Cybernetics, pp. 2008–2013 (2006)
24. Wang, X.Z., Yan, J.H., Wang, R.: A Sample Selection Algorithm in Fuzzy Decision Tree Induction and Its Theoretical Analyses. In: IEEE International Conference on Systems, Man and Cybernetics, pp. 3621–3626 (2007)
25. Baowen, X., Zhou, Y., Lu, H.: An Improved Accuracy Measure for Rough Sets. Journal of Computer and System Sciences 71(2), 163–173 (2005)
26. Elashiry, M.A., Hefny, H.A., Abdel Wahab, A.H.: Induction of Fuzzy Decision Trees Based on Fuzzy Rough Set Techniques. In: Proceedings of the International Conference on Computer Engineering and Systems (ICCES 2011), pp. 134–139. Ain Shams University, Cairo (2011)