Review of Fuzzy Decision Tree: An improved Decision Making Classifier

Vinita A. Gupta
Dept of computer applications, BIT, Durg (C.G.) INDIA
Sunita Soni
Dept of computer applications, BIT, Durg (C.G.) INDIA

ABSTRACT: *Over the years, various methodologies have been investigated and proposed to deal with continuous or numeric data which is very common in any application. With the increasing popularity of fuzzy representation, researchers have proposed to utilize fuzzy logic in decision trees to deal with the situations. This paper presents a survey of current methodology to design FDT (Fuzzy Decision Tree), various issues and applications. The author conclude that fuzzy decision making using decision tree is an emerging technique in terms of applications and there is a enough scope of research in this area.*

Keywords- *Continuous data, Decision tree, Fuzzy logic, Fuzzy decision tree, numeric data,*

I. INTRODUCTION

Data mining, the science and technology of exploring unknown, implicit, potential, novel, useful and intelligible pattern from huge amount of dataset, databases or data warehouse. Data mining is found to be very useful in many areas. Some of them are e-business, banking, medical, data engineering, fuzzy control, construction and maintenance of s/w components etc.

Decision trees are very reliable and efficient tool of decision making which provide high classification accuracy with a simple representation of gathered knowledge and have been used in different areas of decision making. Decision-tree algorithms are found to be one of the most popular methodologies for symbolic knowledge acquisition. Most of the comprehensible decision trees have been designed for perfect symbolic data.

In past years, some new methodologies have been investigated and proposed to deal with uncertain, multi-valued data, and with missing or noisy features. In recent years with the increasing popularity of fuzzy representation, some researchers have proposed to utilize fuzzy representation in decision trees to deal with similar situations and proposed Fuzzy Decision Tree as an improved classifier. Fuzzy decision tree, a new version of ID3 algorithm has been firstly proposed in [1], to generate an understandable classifier using fuzzy sets defined by a user. After that number of methods has been proposed to design fuzzy decision tree and the model has been used in different application areas. In this work we survey the basic characteristics of fuzzy decision trees, different design methods and their applications as a better decision making technique.

II. METHODOLOGY

The major methodology used for this paper was through the survey of journals and publications in the field of medicine, computer science and engineering. The research focused on more recent publications.

III. FUZZY DECISION TREE CLASSIFIER: AN OVERVIEW

3.1 Decision Tree Classifier

A decision tree is a graphic model of a decision making process, and it is generally 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. An example of decision tree has been shown in Fig 1.

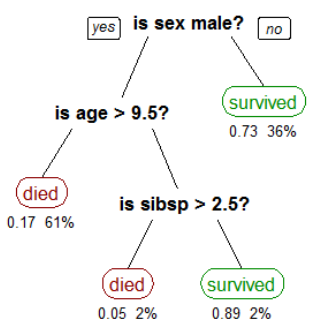

Figure 1.a decision tree from medical domain.

3.2 Fuzzy Logic

Fuzzy logic is a mathematical technique adapted for dealing with uncertain and/or incomplete dataset, and for problems with multiple solutions. It evaluates degrees of truthfulness, as opposed to simply categorizing as completely true or completely false. Instead of only using 0 and 1 as the logic values, it consists of variables that can assume any value from 0 to 1. Fuzzy logic is well adapted for dealing with uncertain and/or incomplete dataset , and for problems with multiple solutions. Software based on application of fuzzy-logic (as compared with that based on Formal Logic) allows computers to mimic human reasoning more closely, so that decisions can be made with incomplete or uncertain data. Fuzzy logic is used for solving problems with expert systems, real-time systems and artificial intelligence applications that must react to an imperfect environment of highly variable, volatile or unpredictable conditions.

3.3 Fuzzy Decision Tree: Issues and Methods

As the decision trees are found to be the most popular choices for learning and reasoning from feature-based examples, a number of alterations have been made in decision tree to deal with language and measurement uncertainties. Classical decision trees work well in crisp domains, but cannot model vagueness. To deal with the crisp boundary problem, fuzzy decision trees have been proposed which combines fuzzy theory with classical trees.

In [12] the authors have introduced a modified version of fuzzy ID3 algorithm that integrates information gain and classification ambiguity to select the test attribute. Decision-tree algorithms are one of the most popular techniques in machine learning. The ID3 algorithm is found to an efficient and popular method for building decision trees that form the basis for many decision tree programs. Fuzzy ID3 is an extension of the existing ID3 algorithm; that integrates fuzzy set theory and ID3 to overcome the effects of spurious precision in the data, to take care of uncertainties in the data and to reduce the decision tree sensitivity to small changes in attribute values. The modified algorithm was found to provide the better accuracy than the original Fuzzy ID3 as well as crisp programs such C4.5. A new machine learning software tool has been introduced based on fuzzy decision trees.

An important modification of combining symbolic decision trees with approximate reasoning offered by fuzzy representation is presented in [3]. The authors have presented a complete method for building the fuzzy tree. Number of inference procedures based on conflict resolution in rule based systems is discussed and an efficient approximate reasoning method is also given. The author has also explored the capabilities of the new framework provided. The resulting learning method is found to be suitable for stationary problems with both numerical and symbolic features, when the goal is to have - high knowledge comprehensibility and gradually changing output.

The fuzzy decision tree is found to be different from traditional decision trees in two respects –

i. Splitting criteria based on fuzzy restrictions, and
ii. An inference procedure.

Fig.3 shows an example of fuzzy decision tree for the data given in Fig. 2.

The author concludes that the fuzzy decision tree can produce real-valued outputs with gradual shifts. Fuzzy sets and approximate reasoning allow for processing of noisy and inconsistent or incomplete data. The effect of such inferior data can be controlled by utilizing various inference methods. These inferences, as well as actual behavior in naturally noisy/incomplete domains, have to be empirically evaluated and reported.

E	*Inc*	*Emp*	*Credit*	*W*
e_1	0.20	0.15	0.0	1.0
e_2	0.35	0.25	0.0	1.0
e_3	0.90	0.20	0.0	1.0
e_4	0.60	0.50	0.0	1.0
e_5	0.90	0.50	1.0	1.0
e_6	0.10	0.85	1.0	1.0
e_7	0.40	0.90	1.0	1.0
e_8	0.85	0.85	1.0	1.0

Figure2. Numeric data.

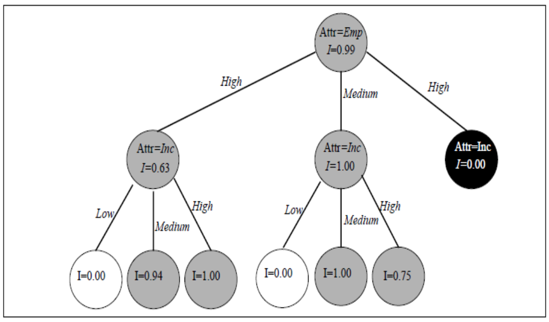

Figure3. Fuzzy Decision Tree

In [4] the author has developed fuzzy decision trees with “Expectation-Maximization algorithm”. In Expectation-Maximization algorithm the database is clustered into groups, unlike the traditional procedures in which the whole database is searched. The proposed method is supposed to overcome the difficulties of the traditional procedures. The major steps followed in the algorithm are-

1. Clustering: The database is clustered into groups as using clustering the searching time is reduced significantly.
2. Fuzzification: The clustered data is fuzzified.
3. Tree building: After fuzzyfying the data, the decision tree is constructed using ID3 algorithm.

The method is applied for the stock market data and adaptation rules are obtained from decision tree constructed**.** To implement the algorithm a case library of numeric real valued was considered. The input training data is having four attributes - Open price, High price, Low price, Close price. The detail steps for classification of the dataset and generation of adaptation rules are as follows-.

i. Clustering case library using E-M algorithm.

The author has attempted to partition the data base into several clusters using E-M algorithm [5]. EM method is simple and easy to implement and it is an iterative method for finding maximum likelihood where the model depends on unobserved latent variables.

1) In this step, the probability of cluster membership of case x_i, for each of the cluster is calculated. Assign each case x_i to cluster C_k with the probability,

$P (x_i \in C_k) = P (C_k/x_i) = P (x_i). P (x_i/C_k)/P$

W

here $P(x_i /C_k) = N(m_k, E_k(x_i))$ follows the normal distribution around mean(m_k) with expectation (E_k).

2) This maximization step computes the distribution parameters and their likelihood given the data. Use the probability estimate from the above to estimate for the model parameters, for example

$$m_k = 1/n\sum_{i=1}^{n} x_i.P(x_i \in C_k)/\sum j P(x_i \in C_j)$$

ii. Fuzzification

Let A_{min} and A_{max} be the maximum and minimum values of an attribute containing N training patterns $\{a_{1h}, a_{2h}, a_{3h},.., a_{Nh}\}$.Triangular-membership function is used for fuzzification and divide A_{min} to A_{max} into three equal intervals. Each attribute has three linguistic terms low, medium, high. The corresponding formulae is given below and graph *r*epresentation is shown in Fig. 4.

Interval = $(A_{max} - A_{min})/3$; $P_{2h} = A_{min}$ + Interval;

$P_{4h} = A_{min} + (2 *$ Interval); $A_{1h} = (A_{min} + P_{2h})/2$;

$A_{2h} = (P_{1h} + P_{4h})/2$; $A_{3h} = (A_{max} + P_{3h})/2)$;

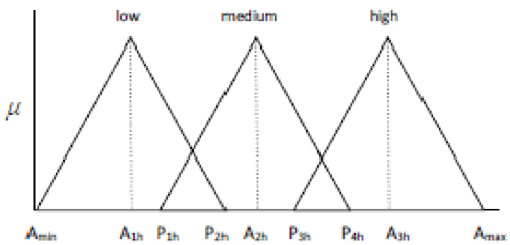

Fig.4 Three membership function

Membership functions for low, medium and high is given below

$$\mu_{low} = \begin{cases} \frac{A_{1h} - a_{ih}}{A_{1h} - A_{min}}, & for\ a_{ih} < A_{1h} \\ \frac{P_{2h} - a_{ih}}{P_{2h} - A_{1h}}, & for\ A_{1h} \le a_{ih} < P_{2h} \end{cases} ;$$

$$\mu_{med} = \begin{cases} \frac{A_{2h} - a_{ih}}{A_{2h} - P_{1h}}, & for\ P_{1h} \le a_{ih} < A_{2h} \\ \frac{P_{4h} - a_{ih}}{P_{4h} - A_{2h}}, & for\ A_{2h} \le a_{ih} < P_{4h} \end{cases} ;$$

$$\mu_{high} = \begin{cases} \frac{A_{3h} - a_{ih}}{A_{3j} - P_{3j}}, & for\ P_{3h} \le a_{ih} < A_{3h} \\ \frac{A_{max} - a_{ih}}{A_{max} - A_{1h}}, & for\ A_{3h} \le a_{ih} < A_{max} \end{cases} .$$

iii. Generating decision tree using ID3 learning algorithm:

ID3 uses an information –theoretic approach. The procedure is that at any point, one examines the feature that provides the greatest gain in information or, equivalently, the greatest decrease in entropy.

The general case is that of N labled patterns partitioned into sets of patterns belonging to classes C_i, i = 1,2,3….,*l*. The population in class C_i is n_i. Each pattern has n features and each feature can take on two or more values. The ID3follows the steps as below-

1) Calculate initial value of entropy,

$$Info(D) = -\sum_{i=1}^{m} p_i \log_2(p_i), \text{where } p_i = (n_i/N)$$

2) Select that feature which results in the maximum decrease in entropy or

maximum gain in information, to serve as the root node of the decision tree

$$Gain(A) = Info(D) - Info_A(D)$$

where $Info_A(D)$ is expected information required to classify the tuple from D, based on the partitioning by attribute A.

3)Build the next level of the decision tree providing the greatest decrease in entropy.

4) Repeat step 1 through step 3 and continue the procedure until there are no attributes for more classification.
At this stage, a set of leaf nodes of the decision tree were obtained.

The example decision tree obtained for single cluster from the above procedure for the stock exchange data is given below

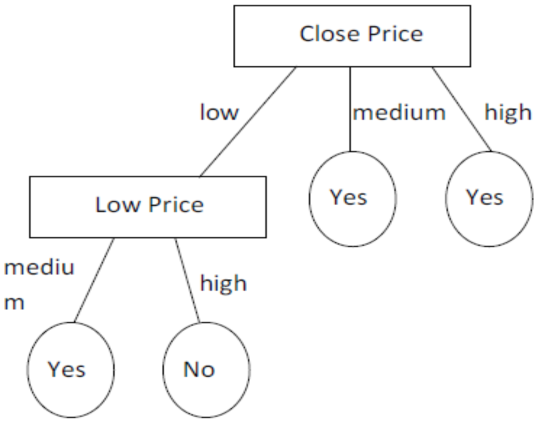

Figure.5 Decision tree generated for a single cluster

iv. Generating adaptation rules from decision tree

After generating the decision tree by using ID3 algorithm, a set of adaptation rules can be extracted from the tree.
Each path from the root to a leaf is expressed as an adaption rule. Thus the number of leafs is just the number of adaption rules. For example from decision tree shown in Fig. 4, The general form of adaptation rule generated from the decision tree for the maximum number of antecedents for each fuzzy rule which is limited to two, is as follows,
Rule: IF price of X1 is [Small|Medium|High] AND price of X2 is [Small|Medium|High] THEN the action=[Yes|No],
Where X = {Open Price, High price, Low price, Close price}.

The decision tree is executed by starting from the root node and repeating to test the attribute at the node and branch to an edge by its value until reaching at a leaf node, a class attached to the leaf being as the result

The author concludes that the proposed methodology can be used to predict the future action for a given tuple using the existing data. This model can applied on any numerical huge database. For example, the model can be applied on medical datasets related to liver disorder, heart problems, kidney failure, nervous system response etc., can also be applied to predict the future increase or decrease values in the business. For example to predict the daily increase or decrease in the gold prices or silver prices in the market.

In this paper [5], author has presented a new method of fuzzy decision trees called soft decision trees. The proposed method combines tree growing and pruning, to determine the structure of the soft decision tree, with refitting and back fitting. The refitting and back fitting improve its generalization capabilities. The proposed method is explained and motivated and is first analyzed empirically on 3 large databases in terms of classification error rate, model complexity and CPU time. This model is studied on 11 standard UCI Repository databases which shows that the soft decision trees produced by the proposed method are significantly more accurate than standard decision trees.

In [9] authors conclude that compared to other fuzzy models, simple fuzzy logic rules (IF ... THEN... rules) based on triangular or trapezoidal shape fuzzy sets are much simpler and easier to understand. Further choosing the right combination of attributes and fuzzy sets which have the most information is the key point to obtain good accuracy in fuzzy rule based learning algorithms.

In [13] the authors analyze some promising variants of building classification rules from a given fuzzy decision based on cumulative information. Integrating fuzzy logic algorithms into databases allows reducing uncertainty of data in the databases and to increase discovered knowledge's accuracy. The work done compared the classification accuracy with the accuracy which is reached by statistical methods and other fuzzy classification rules. The main purpose of our experimental study is to compare the two ways of using fuzzy classification rules made from FDT based on cumulative information for classification. Their classification accuracy is also compared with other methods.
The experiments have carried out on Machine Learning databases. First, the databases were fuzzyfied. Around 70% of the databases have been used for building classification models. The remaining 30% have been used for verification of the models. The error rate is calculated as the ratio of the number of misclassification combinations to the total number of combinations. The

experimental results are in shown in Table.3, where CA-MM denotes FDT-based fuzzy classification rules algorithm, CA-SP denotes a CA-MM's modification, CI-RM-M denotes for making FDT with cumulative information and which uses the second way of fuzzy classification rules induction, CI-RM-0 uses the first way of fuzzy classification rules induction, NBC denotes Naïve Bayes Classifier, k-NN denotes k-Nearest Neighbour Classifier. The last row contains average error rate for all databases and respective methods.

Table1. Comparison of various classifications Algorithm

Database	Error rate for given database and classification method.					
	CA-MM	CA-SP	CI-RM-M	CI-RM-O	NBC	k-NN
BUPA	0.4253(4)	0.4253(4)	0.4174(2)	0.4205(3)	0.4416(6)	0.3910(1)
Ecoli	0.3112(6)	0.3053(5)	0.2022(2)	0.2654(4)	0.1547(1)	0.2046(3)
Glass	0.4430(4)	0.4294(3)	0.3988(2)	0.4647(5)	0.5394(6)	0.3335(1)
Haberman	0.2615(3)	0.2615(3)	0.2624(5)	0.2609(2)	0.2453(1)	0.3471(6)
Iris	0.0400(2)	0.0400(2)	0.04067(4)	0.02956(1)	0.04556(5)	0.05044(6)
Pima	0.2509(4)	0.2460(2)	0.2436(1)	0.2563(5)	0.2483(3)	0.3091(6)
Wine	0.08170(6)	0.06094(4)	0.04566(2)	0.06509(5)	0.02697(1)	0.05094(3)
Average	**0.2951**(6)	**0.2526**(5)	**0.2301**(1)	**0.2518**(4)	**0.2431**(3)	**0.2409**(2)

V. APPLICATIONS OF FUZZY DECISION TREE

The Fuzzy Decision Tree is found to be an efficient alternative to crisp classifiers that are applied independently. The cooperation of Fuzzy Logic and decision trees tries to soften the accuracy/interpretability tradeoff. Thus induction of FDT is very 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 [2].

In paper [6], authors have shared their experiences of investigating intelligent machine learning techniques for breast cancer prognosis analysis. They have analyzed the possible potential of fuzzy logic based classifiers, and concluded that they are fit to act as natural allies of a physician involved in predictive medicine. Authors have outlined some future dimensions which can help wFDTs to prove their potential as a strong classifier and predictor in cancer prognosis.

In [10] Fuzzy decision tree has been used to extract linguistic rules to develop a fuzzy knowledge-based network. A scheme is formulated for automatic linguistic discretization of continuous attributes, based on quintiles. A new concept has been developed to measure the goodness of a decision tree, in terms of its compactness (size) and efficient performance. Linguistic rules are mapped to a fuzzy knowledge-based network, having the frequency of samples and depth of the attributes in the decision tree. The effectiveness of the system is evaluated using three sets of real-life data in terms of various parameter namely recognition scores, structure of decision tree, performance of rules, and network size.

In [11] authors concluded that Fuzzy Decision Tree is becoming more and more significant areas of application in different platforms namely medical, educational, chemical and multimedia. The work done investigated the applications of fuzzy decision tree in heterogeneous fields in real life. The major areas where the decision tree has been successfully applied are intrusion detection, flexible querying (modus ponens), analysis of cognitive process (Human Computer Interaction), for user authentication in biometrics, as parallel processing support, in stock-market, for information retrieval and data mining etc. The powerful combinatorial methods found in fuzzy logic have been used to prove fundamental results in various application areas.

VI. CONCLUSION

This paper presents a survey of current methods for FDT (Fuzzy Decision Tree) designs, various existing issues and its applications. The survey concludes that due to potential advantage of Fuzzy decision tree over the traditional decision tree it has been applied in many real-world applications for better performance ranging from intrusion detection, flexible querying (modus ponens), analysis of cognitive process (Human Computer Interaction), for user authentication in biometrics, as parallel processing support, in stock-market, for information retrieval and data mining. The author feels that still there is lot of scope of work in different application areas, like in the field of medical decision making where quantitative attributes are discritized to get transformed binary database. In such data base each record fully belongs to only one fuzzy set and suffers the crisp boundary problem. Fuzzy logic can be used to deal the crisp boundary problem which will ultimately result the improved accuracy in decision making.

REFERENCES

1. Umano, M., Okamoto, H., Hatono, I., Tamura, H., Kawachi, F., Umedzu, S. Kinoshita, J Fuzzy decision trees by fuzzy ID3 algorithm and its application to diagnosis systems, Fuzzy Systems, IEEE World Congress on Computational Intelligence., Proceedings of the Third IEEE Conference on 1994.
2. Tao Wang, Zhoujun Li, Yuejin Yan, Huowang Chen, A Survey of Fuzzy Decision Tree Classifier Methodology, Fuzzy Information and Engineering Advances in Soft Computing Volume 40, 2007, pp 959-968.
3. Cezary Z. Janikow, Fuzzy Decision Trees: Issues and Methods , Fuzzy Decision Trees: Issues and Methods.

4. Daveedu Raju Adidela, Jaya Suma. , Lavanya Devi. G, Construction of Fuzzy Decision Tree using Expectation Maximization Algorithm, International Journal of Computer Science and Management Research Vol 1 Issue 3 October 2012, ISSN 2278-733X.

5. A.P. Dempster, N.M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," Journal of the Royal Statistical Society, Series B, vol. 39, no. 1, pp. 1-38, 1977.

6. Umer Khan, Hyunjung Shin, Jong Pill Choi, Minkoo Kim, wFDT - Weighted Fuzzy Decision Trees for Prognosis of Breast Cancer Survivability

7. Vitaly Levashenko, Elena Zaitseva, Fuzzy Decision Trees in Medical Decision Making Support System, Proceedings of the Federated Conference on Computer Science and Information Systems pp. 213–219

8. Khan MU, Predicting breast cancer survivability using fuzzy decision trees for personalized healthcare, , Conf Proc IEEE Eng Med Biol Soc. 2008;2008:5148-51.

9. J. F. Baldwin, Dong (Walter) Xie, Simple Fuzzy Logic Rules Based on Fuzzy Decision Tree for Classification and Prediction Problem**,** Intelligent Information Processing II, Volume 163, 2005, pp 175-184.

10. Sushmita Mitra, , Kishori M. Konwar, and Sankar K. Pal, Fuzzy Decision Tree, Linguistic Rules and Fuzzy Knowledge-Based Network: Generation and Evaluation , IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART C: APPLICATIONS AND REVIEWS, VOL. 32, NO. 4, NOVEMBER 2002.

11. Kavita Sachdeva, Madasu Hanmandlu and Amioy Kumar, Real Life Applications of Fuzzy Decision Tree. *International Journal of Computer Applications* 42(10):24-28, March 2012. Published by Foundation of Computer Science, New York, USA.

12. Abu-halaweh, N.M, Harrison, R.W., Practical fuzzy decision trees, Computational Intelligence and Data Mining, 2009. CIDM '09. IEEE Symposium on Page(s):211 – 216, 2009.

13. Karol MATIAŠKO, Ján BOHÁČIK, Vitaly LEVASHENKO, Štefan KOVALÍK, Learning Fuzzy Rules From Fuzzy Decision Trees, Journal of Information, Control and Management Systems, Vol. 4, (2006), No. 2 143.