Induction of Fuzzy Pattern System

Shruti S. Pophale
PG student - Department of Computer Science and Engineering
MIT , Aurangabad, Maharashtra, India

S. A. Kinariwala
AP-Department of Computer Science and Engineering
MIT , Aurangabad, Maharashtra, India

Abstract - Machine learning, data mining, and several related research areas are concerned with methods for the automated induction of models and the extraction of patterns from realistic data. Fuzzy classification is one of the important applications in fuzzy set where values range between 0 and 1. Main objectives of the fuzzy logic and fuzzy set to find a set of fuzzy rules that form a classification model. The set is major advantage of using fuzzy rules for classification applications to maintain transparency as well as high accuracy rate. Fuzzy pattern tree induction introduced as a novel for machine learning method for classification. Pattern tree is a tree-like structure, whose inner nodes are marked with fuzzy logical operators or arithmetical operators and whose leaf nodes are associated with fuzzy terms on input attributes. Pattern tree make use of different aggregation including t-norms and t-conorms. There are two types of Fuzzy Pattern trees: (i) Bottom up induction: Information propagates from bottom to top. (ii) Top down induction: Information propagates from top to down. In this paper performance analysis and comparative study of above both approaches using number of parameters like accuracy rate, model size is mentioned. The paper is divided into five sections: section 1 is introduction of Fuzzy pattern tree, section 2 describes types of fuzzy pattern tree schemes and section 3 presents their comparative performance, section 4 Results, section 5 does the conclusion.

I. INTRODUCTION

A. Fuzzy set and Fuzzy rules:

Fuzzy set is an extension of a crisp set. Crisp set allows value like 0 and 1 or yes and no. It is defined in eq. (1). Fuzzy set allows value between 0 and 1 or value between yes and no. It is defined in eq.(2).A Crisp sets allow only full membership or no membership at all and fuzzy sets allow partial membership. Besides, each element may belong to more than one set [4]. In a crisp set, the membership of an element x in set A is described by a characteristic function $\mu_A(x)$ where

$$\mu_A(x) = \begin{cases} 1 & if\ x \in A \\ 0 & if\ x \notin A \end{cases} \qquad (1)$$

The fuzzy-set extends this concept by defining partial memberships, which can take values ranging from 0 to 1. A membership function to convert crisp value into fuzzy value is formally defined as follows

$$\mu_A : X \rightarrow [0,1] \qquad (2)$$

where X refers to the universal set for a specific problem.

B. Fuzzy Pattern tree:

Fuzzy pattern tree induction introduced as a novel for machine learning method to build classification model. Pattern tree is a hierarchical tree-like structure, whose inner nodes are marked with fuzzy logical operators or arithmetical operators and whose leaf nodes are associated with fuzzy values on input attributes. The fuzzy terms of input variables are on at different levels except the root of the tree. It uses fuzzy aggregations to aggregate from bottom level to the top level that is root node and also top level to the bottom that is leaf node. A classification application requires many classes which involve several outputs as the classes. Classification model should have many pattern trees as the number of output classes with every pattern tree representing one class. It implements a recursive function that maps a combination of attribute values, entered in the leaf nodes, to a number in the unit interval, produced as an output by the root of the tree. Fuzzy aggregations are fuzzy logical and arithmetic operators. Already due to the use of limited number of nonparametric operators in fuzzy pattern trees, a wide variety of operators for aggregation of fuzzy predicates actually was ignored and this could have the impact on the performance and results of fuzzy pattern trees. The arithmetic operators can be weighted arithmetic mean and ordered weighted arithmetic mean [3].These operators are parameterized, which means that implicitly select from an infinite number of operators

but the logical operators, t-norms and t-conorms ,as describes in above section, used in fuzzy pattern trees are not parameterized. There are several fuzzy aggregation families which contain the standard t-norms (t-conorms) as special cases [5].
From a modeling point of view, the pattern tree approach is based on three important conceptions [6]:

- Fuzzification of input attributes.
- Hierarchical structuring of a functional dependency.
- Flexible aggregation of sub-criteria by means of fuzzy operators.

If data is not in fuzzy form then fuzzification is needed. It can be done by using membership functions. Both types of fuzzy pattern trees require recursive process so hierarchical structuring using fuzzy operators is also required.

C. Fuzzy Operators:

Pattern tree contains inner operators which are fuzzy operators & aggregation operators.

(I) t-norms and t-conorms:

Fuzzy aggregations are logical operators applied to fuzzy membership values or fuzzy sets. They have three subcategories, namely t-norm that is triangular norms and
t-conorm that is triangular conorms and averaging operators such as weighted averaging (WA) and ordered-WA (OWA).t-norms were introduced by Schweizer and Sklar to model distances in probabilistic metric spaces. In fuzzy sets theory, t-norms and t-conorms are broadly used to form logical operators AND and OR. A *t*-norm is a generalized conjunction, namely, a monotone, associative, and commutative mapping with neutral element 1 and absorbing element 0. Likewise, a *t*-conorm is a generalized disjunction, namely, a monotone, associative ,and commutative mapping with neutral element 0 and absorbing element 1. The basic t-norm and t-conorm pairs that operate on two fuzzy membership values a and b, where a, b ϵ [0, 1], that means a, b are fuzzy values which is shown in following tables: Table 1 and Table 2 shows t-norms and t-conorms respectively [1].

Table 1 Fuzzy Operators: t-norms

Name	**Definition**	**Code**
Minimum	min{a, b}	MIN
Algebraic	Ab	ALG
Lukasiewicz	max{a+b-1,0}	LUK
Einstein	$\frac{ab}{2-(a+b-ab)}$	EIN

Table 2 Fuzzy Operators: t-conorms

Name	**Definition**	**Code**
Maximum	max{a, b}	MAX
Algebraic	a+b-ab	COALG
Lukasiewicz	min{a+b-1,1}	COLUK
Einstein	$\frac{a+b}{1+ab}$	COEIN

(II) WA and OWA:

A weighted average that is WA[7] operator of dimension n is mapping $R^n \rightarrow R$ that has an associated n-element vector
$w=(w_1,w_2,\ldots\ldots,w_n)^T$,$w_i \in [0,1]$, $1 \leq i \leq n$ and
$\sum_{j=1}^{n} w_i=1$, so that

$$WA(a_1, \ldots, a_n) = \sum_{j=1}^{n} w_j a_j \qquad (3)$$

An Ordered Weighted Average OWA operator [8] of dimension n is a mapping $R^n \rightarrow R$ that has an associated n-element vector $w=(w_1,w_2,\ldots\ldots,w_n)^T$, $w_i \in [0,1]$, $1 \leq i \leq n$ and $\sum_{j=1}^{n} w_i=1$, so that

$$OWA(a_1, \ldots, a_n) = \sum_{j=1}^{n} (w_j f_j(a_1, \ldots, a_n)) \quad (4)$$

Where $f_j(a_1, \ldots, a_n)$ returns j^{th} biggest element of the collection $(a_1, \ldots, a_n)$.

II) APPLICATIONS OF PATTERN TREE:

Fuzzy Pattern trees are used for regression and fuzzy system modeling [6].This system proposed a variant of pattern tree induction suitable for learning real-valued functions. Fuzzy pattern trees are also used in Linguistic data mining with fuzzy FP-trees [4]. Due to the increasing occurrence of very large databases, mining is very useful in extracting information and knowledge from transactions which involves an important research area. Many algorithms already exists for mining most of which were based on items with binary values. Transactions with quantitative values are, commonly used in real-world applications. So system requires the frequent fuzzy pattern tree for extracting frequent fuzzy item sets from the transactions with quantitative values. When extending the FP-tree to handle fuzzy data, the processing becomes much more complex than the original since fuzzy intersection in each transaction has to be handled. The fuzzy pattern tree construction algorithm is thus designed, developed and the mining process based on the tree is presented. Fuzzy pattern tree used for Binary classification on data stream[9].An evolving version of the fuzzy pattern tree classifier that meets the increased requirements of incremental learning on data streams. The main objective is to maintain, with the current model, a set of neighbor trees that can replace the current model if the performance of the latter is no longer optimal. Thus, a modification of the current model is required implicitly in the form of a replacement by an alternative tree that is fuzzy pattern tree. A replacement can be made on the basis of the performance of all models. Learning Pattern tree classifiers using a Co-Evolutionary algorithm is proposed [10]. Main idea is it is an alternative method for learning pattern tree classifiers. This work was mainly motivated by two alleged disadvantages of the original learning method: First, it is based on a simple greedy method and, therefore, search strategy, which is likely to get stuck in local optima. Second, it seems to solve a problem which is actually more difficult than necessary. This disadvantage can be overcome with Co-evolutionary algorithm which is more sophisticated search method. Moreover, instead of maximizing a similarity function for each class separately, the interdependency of the individual pattern trees is taken into account and seeks to maximize classification accuracy directly. Fuzzy pattern trees are also very useful in marketing fields and medical fields to make predication and to make classification.

III) TYPES OF FUZZY PATTERN TREE:

A. Bottom Up approach:

As the name suggest in bottom up approach in pattern tree information propagates from bottom to top. It is proposed in [2]. A node takes the values of its descendants as input and combines them using the respective operator, and then submits the output to its predecessor. Thus, a pattern tree implements a recursive mapping producing outputs in the unit interval. A pattern-tree classifier consist group of pattern trees: one for each class. A query instance which is to be classified is submitted to each tree, and a prediction is made in favor of the class whose tree produces the highest output. For example to check income of person is greater than or equal to 50K/yr or not we can draw the tree by following way:

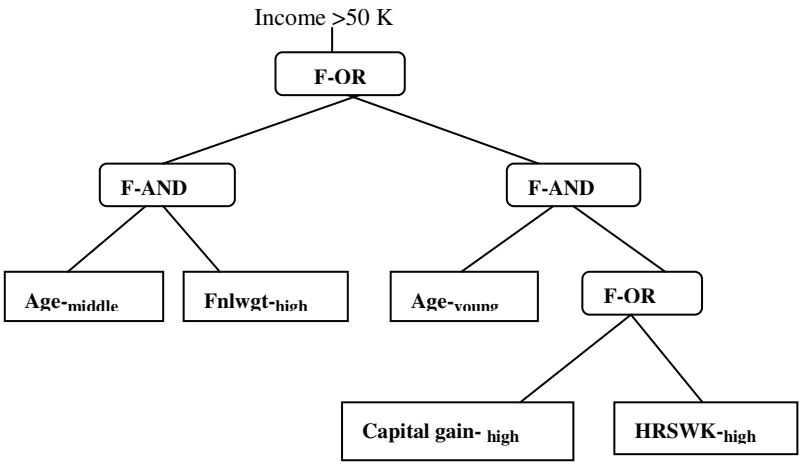

Figure.1 Pattern-tree Adult Income example

This tree can be read as person can have income is greater than 50k/yr who is either middle aged and financial weight (Fnlwgt) is high or capital gain is high and hours per week worked is high or age is young. A tree of this kind could have been induced from a dataset containing descriptions of adult in terms of attributes like age, financial weight, hours worked per week etc., as well as information about whether income is greater than 50k/yr or not.
Attributes are all in fuzzy like age is middle or young, capital gain is high or low etc .Rounded Rectangle represents fuzzy operators and square represents attributes in data set with fuzzy values. Fuzzy pattern tree can be represented for wine-quality example in fig no 2:

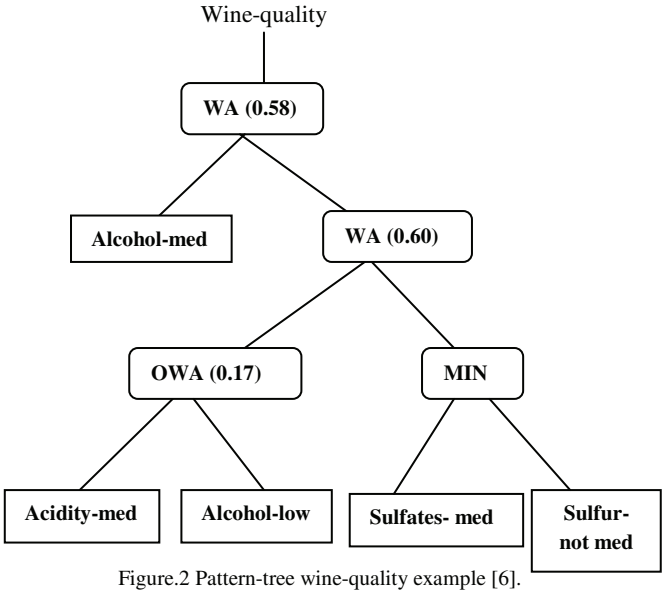

Figure.2 Pattern-tree wine-quality example [6].

Referred tree has root class wine-quality and leaf is with attribute and with linguistic terms like high, low, med etc. The tree can be interpreted in a quite intuitive way: The wine quality is an average of two criteria, namely a medium level of alcohol and a second criterion. This second criterion is in turn a complex one, namely an average of two sub-criteria. One of these sub-criteria is the conjunctive condition that the level of sulfates is medium and the level of sulfur not medium, and the second sub-criterion can be interpreted correspondingly[6].
Pattern tree by bottom up approach is built one by one. For each class, the induction method performs as follows [2]:
1) Initialize with primitive pattern trees.
2) Select candidate's trees by evaluation of their similarity to the target class.
3) Check stopping criterion.
4) Recombine candidates using fuzzy operators;.
5) Go to step 2 till stopping criteria is not satisfy
In this way pattern tree can build by bottom up approach.

B. Top-down approach:

Basic idea of top down approach is represented in [1]. In this approach, one population of pattern trees is maintained for each class, and the development of these populations is coordinated by means of a global fitness function. Top down tree is different from bottom up tree by the direction and also construction of bottom up tree done always by merging two candidate trees into a completely new tree. In top down approach tree construction is done by expanding one of its leaf nodes that is by replacing with a three- node tree. Example can be as follows:

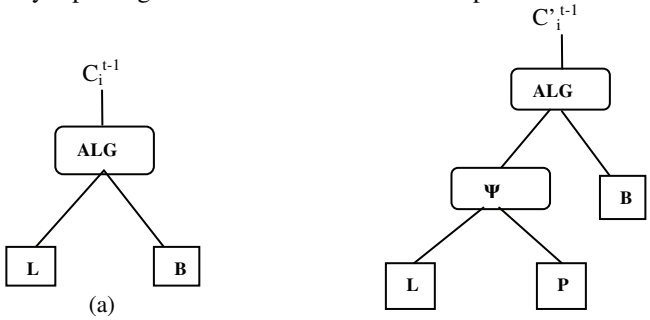

Fig.3.Top-down induction. (b)

Fig 3(a) shows pattern tree at the iteration C_i^{t-1}

Fig 3(b) at the iteration $C'_i{}^{t-1}$ tree is expanding by replacing L node by the new node L and P where **Ψ** is any fuzzy operator that is arithmetic or logical.

Top down tree can work by following way [1]:

1) Initialize with primitive tree.
2) Candidate trees are initialized by the best primitive trees.
3) New candidate trees are created by replacing exactly one leaf node.
4) Try for all replacement of all leaf nodes of candidate tree.
5) The new candidate tree thus obtained is then evaluated by computing its similarity to the target class.
6) Check for termination criteria.
7) Best candidates are selected and passed to the next iteration, unless the termination criterion is fulfilled.

IV) HANDLING OF MISSING VALUES:

In this system missing values are handles in data preprocessing. K-nearest neighbor algorithm is used to find missing values. It identifies nearest value for missing data. It creates first training files for all attributes .This training files created without missing values .This are input for finding missing values. In this training files fuzzy terms are used like F0,F1…Fn and according to that ranking is done. Then it's nearest fuzzy term is assigned to missing value and actual value is computed as follows where x is missing value and max, min is range for selected fuzzy term.

$$x=\frac{max-min}{2}$$

V) PERFORMANCE EVALUATION OF SYSTEM

Fuzzy Pattern trees are compared with various parameters.

A) Accuracy Rate: It is calculated for classification of data. So Accuracy is calculated by using formula:

$$Accuracy\ rate = \frac{Number\ of\ classes(actual)predicted}{Number\ \ of\ to\ be\ predicted}$$

B) Runtime**:** It gives total time required to test the data. It is the difference between times of start the testing up to end of total data to be tested. It is measured in milliseconds.

C) Model Size**:** The Model size is given by the average number of nodes contained in one fuzzy pattern tree.

D) Similarity Measure: It is calculated to find similarity from the class. That is how predicted output class is similar to the target class. Formula is given both types algorithm. It is calculated in %. It is plotted against iteration, iteration on X-axis and similarity value on Y-axis.

Following table represents results with above defined factors for iris-missing values.

Table 3 Results

Parameters	30 % Dataset for training				80 % Dataset for training			
	PTTD.1	PTTD.5	PTTD.25	Bottom up	PTTD.1	PTTD.5	PTTD.25	Bottom up
Accuracy Rate	0.866666	0.866666	0.866666	0.809523	0.86666	0.86666	0.86666	0.86666
Runtime	78	47	78	109	406	156	468	437
Model Size	9	27	17	29	23	21	29	32
Similarity Measure	0.271525	0.269165	0.272286	0.26146	0.304971	0.3047744	0.304971	0.309127

VI. CONCLUSION:

Fuzzy pattern trees are very useful in many applications like medical fields, marketing fields and also to make classification model. In predictive applications also fuzzy pattern trees are very useful. In both approaches defined in

above section dataset taken without missing values. So to handle such situation new techniques k-nn used. Runtime is more when k-nn is used and accuracy is good.

REFERENCES

[1] Robin Senge and E. H¨ullermeier, "Top down induction of fuzzy pattern trees ",IEEE TRANSACTIONS ON FUZZY SYSTEMS, VOL. 19, NO. 2, APRIL 2011.

[2] Z. Huang, T. D. Gedeon, and M. Nikravesh, "Pattern trees induction: A new machine learning method," IEEE Trans. Fuzzy Syst., vol. 16, no. 4,pp. 958–970, Aug. 2008.

[3] S. Rajaeipour, G. Shojatalab " Uni-norm Fuzzy Pattern Trees for Evolving Classification by Imperialist Competitive Algorithm", The Journal of Mathematics and Computer Science Vol. 4 No.3 (2012) 502 - 513

[4] Chun-Wei Lin, Tzung-Pei Hong, Wen-Hsiang Lu," Linguistic data mining with fuzzy FP-trees" Science direct, Expert Systems with Applications 37 (2010) 4560–4567

[5] E. P. Klement, R. Mesiar, and E. Pap, Triangular Norms. Norwell, MA:Kluwer, 2002.

[6] Robin Senge and E. H¨ullermeier, "Pattern trees Regression and Fuzzy Systems Modeling ",Proc.WCCI- 2010

[7] B. Schweizer and A. Sklar, Probabilistic Metric Spaces. NewYork:North-Holland, 1983.

[8] R. Yager, "On ordered weighted averaging aggregation operators in multi criteria decision making," IEEE Trans. Syst., Man Cybern., vol. 18, no. 1,pp. 183–190, Jan./Feb. 1988.

[9] R. Senge and E. H¨ullermeier, "Learning pattern tree classifiers using a co-evolutionary algorithm," in Proc.19th Workshop Comput. Intell.,F. Hoffmann and E. H¨ullermeier, Eds. Dortmund, Germany: KIT Sci.,2009, pp. 22–33.

[10] Ammar Shaker, Robin Senge and E. H¨ullermeier, "Evolving fuzzy pattern trees for binary classification on data streams".

[11] video