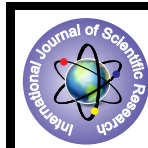


Decision Trees in Fuzzy Logic



Mathematics

KEYWORDS: words:Construction of Decision Trees,Applications

VINNAKOTA MANJULA

Basic Engineering Department,
DVR& Dr. HS MIC College of Technology, Kanchikacherla

ABSTRACT

This paper describes the Application of Decision Trees in Fuzzy Logic. Decision trees are one of the most popular choices for learning and reasoning Decision Tree analysis involves the construction of diagram that shows at a glance when decisions are expected to be made- in what sequence, the possible consequences etc. This paper, describes an introduction to fuzzy decision trees. They have undergone a number of alterations to deal with language and measurement uncertainties combined by symbolic decision trees with approximate reasoning offered by fuzzy representation. The intent is to exploit complementary advantages of both popularity in applications, and the ability to deal with inexact and uncertain information of fuzzy representation.. In particular, knowledge inferences must be newly defined for the fuzzy tree. Decision-tree algorithms are of the most popular applications in real life The learning method is most suitable for stationary problem.

Introduction

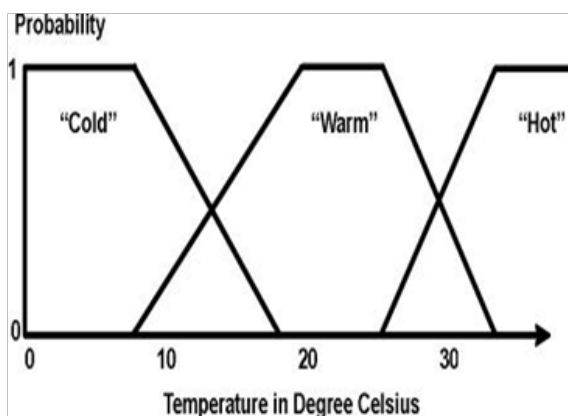
In almost every real-life every one is confronted with growing amounts of data coming from measurements, or from manual data registration, centralization procedures, and, most often, it would be a waste not to take advantage of these data. Recent developments in data storage devices, database management systems, computer technologies and automatic learning techniques make data analysis tasks easier and more efficient. In this context, data mining is a modern concept beginning to be widely used. The general purpose of data mining is to process the information embedded in data so as to develop better ways to handle data and support future decision-making. Machine learning, association rules, clustering methods, artificial neural networks, statistical and visualization tools are common techniques used in data mining. Decision tree techniques are able to treat large scale applications. But they are also recognized as highly unstable classifiers with respect to minor perturbations in the training data. Fuzzy logic brings in an improvement in these aspects due to the elasticity of fuzzy sets formalism

Construction

A fuzzy decision tree can be constructed by allowing the possibility of partial membership of a point in the nodes that make up the tree structure. This extension of its expressive capabilities transforms the decision tree into a powerful functional approximant that incorporates features of connectionist methods.. Fuzzification is achieved by superimposing a fuzzy structure over the skeleton of a CART decision tree. A training rule for fuzzy trees, similar to backpropagation in neural networks. This rule corresponds to an optimization algorithm that fixes the parameters of the fuzzy splits. The method developed for the automatic generation of fuzzy decision trees and it will be applied to both classification and regression problems. In regression problems, the continuity constraint imposed by the function representation of the fuzzy tree leads to substantial improvements in the quality of the regression and limits the tendency to overfitting. In classification, fuzzification provides a means of uncovering the structure of the probability distribution for the classification errors in attribute space. This allows the identification of regions for which the error rate of the tree is significantly lower than the average error rate.

A geographer examines a thermal satellite image of a city, where the surface temperatures can be derived from the different colours of the pixels. The geographer wants to distinguish "cold", "warm" and "hot" areas of the city for a bio climatological model.

Figure-1



Most commonly an analysis will then use the fuzzy set and design logical rules. Therefore, fuzzy logic is rule based. A rule indicates a feature in the thermal picture is a heated house if it is more than 0,6 of "warm" in the day and more than 0,4 of "warm" during the night.. A customer wants to know if a property fits or not or if something is inside a class or outside..

Crisp regression tree

The regression tree has test nodes and terminal nodes. Each node box is stamped with the local estimation of the output. Under each test node, the selected test appears as a condition on a single attribute at a time, regarding a single threshold and having two possible answers: yes or no The local input space is thus split into two non-overlapping subregions of objects. The objects in one such sub region should ideally have the same output value. The tree may be translated into an equivalent set of mutually exclusive rules, each one corresponding to a path from the top node to a terminal node.

Soft decision tree

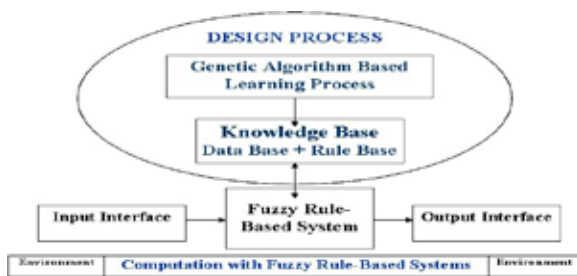
The soft decision tree has also test nodes and terminal nodes. Each node is also marked with its local estimation of the output. Under each test node, the selected test appears as a condition on a single attribute at a time, regarding a pair of two parameters values in brackets. These two parameters characterize the function called discriminator needed to fuzzily split the local set of objects of a given current test node. A widely used shape of discriminator function is the piecewise linear one. The two parameters are location of the cut-point and corresponds to the split threshold in a test node of a decision or a regression tree, and which is the. With such a piecewise linear discriminator

function, the local input space of a node is split into two overlapping sub regions of objects. The larger the transition region in a test node, the larger the overlap and the softer the decision in that node. in the simplest case through one path, in the most complex case through all the paths. Thus, the node may be seen as a fuzzy set.

Applications

Decision-tree algorithms are one of the most popular applications in machine learning. The ID3 algorithm is an efficient method for building decision trees that form the basis for many decision tree programs. Fuzzy ID3 is an extension of the existing ID3 algorithm; it integrates fuzzy set theory and to treat uncertainties in the data and to reduce the decision tree sensitivity to small changes in attribute values. In many real world applications the data used are imprecise and subjective nature. When an expert tries to analyze a certain event, or what ever in a real world environment, they will express the estimates with some degree of confidence. These estimates can be represented as numerical or logical values. The use of numerical values makes the search of common characteristics from the input attributes more difficult because of variety of real values.

Figure-2



>Intrusion Detection

Developing effective methods for the detection of intrusions and misuses is essential for assuring system security. Various approaches to intrusion detection are currently being in use with each one has its own merits and demerits. The objective of this study is to test and improve the performance of a new class of decision tree. The C-fuzzy decision trees are classification constructs that are built on a basis of information granules fuzzy clusters. The way in which these trees are constructed deals with successive refinements of the clusters forming the nodes of the tree. When growing the tree, the nodes will split into granules of lower diversity. The performance, robustness, and usefulness of classification algorithms are improved when relatively few features are involved in the classification. The purpose is to identify that is computationally efficient and effective. The focus is on improving the performance by reducing the number of features and selecting more appropriate data set.

>Flexible Querying

Nowadays, a large amount of data is contained in a lot of databases. These data represent an important source of potential knowledge to use. Data mining is the process of mining databases in order to induce such knowledge. Fuzzy data mining is concerned by the mining of fuzzy knowledge. It enables us to improve the querying process of the database. A set of rules can be introduced to optimize the query process of the database or to deduce decisions from data. Such a set of induced fuzzy rules can be associated with a database as a knowledge base that can be used to help answering frequent queries. A fast response can be found for a query on the value of an attribute. It can also lower the conditions on the values of attributes for a query, before the process of examination of the database.

- Moreover, fuzzy rules can take advantage of the fuzziness of their values to take into account new numerical or fuzzy values. The method of classification with such a set of fuzzy rules is a good way to handle new values for attributes.

- Secondly, a set of fuzzy rules is completely understandable and a decision taken by means of these rules is explainable. It can be used as a new knowledge on the domain of the database, and it can be understood by any expert of this domain. The method of classification by means of a fuzzy decision tree lies on the use of measure of satisfiability to compare observed values to testing values.

- Moreover, by using a fuzzy decision tree instead of a classical decision tree, a slight change in the values of a description leads to a slight change in the value of the decision. Given a measure of satisfiability and a process of inference by means of a fuzzy decision tree, shows that a continuity in the value of the decision is obtained relatively to the values of the description.

>User Modeling from Human- Computer Interaction

A fundamental objective of human-computer interaction research is to make systems more usable, more useful, and to provide users with experiences fitting their specific background knowledge and objectives.

For Providing Advices system uses a learning agent, a fuzzy decision tree, which works on traces of human-computer interactions by providing advices that come from people who have the same cognitive behavior. The TAFPA software is based as an analysis of the cognitive activity of a user is completed to provide hints...

>User Authentication in Biometrics

Biometric based personal authentication can be viewed as procedural technique where the claimed identity is authenticated based on user's physiological or behavioral traits. The principal steps in authentication system generally include

the acquisition of the biometric trait, feature extraction, generation of matching scores, and calculation of error rates. The best choice in a system is to plot the two associated error rated curve and select the suitable measure according to the tradeoff between the error rates. To design a classification system which can achieve such accuracy, two very important considerations are

- 1.The choice of a suitable classifier with relevant classification technique
- 2.The choice of biometric features which are used for training the classifier.

The neural network based classifier is the most frequently used classification system utilized by several authentication systems. It can be used to generate decision to get the final trained tree. Each data has a fuzzy representation according to the membership function and it can belong to a particular class based on this representation. The final decision about the classification is to make a real time biometric authentication security system using fuzzy decision tree as a supported classifier.

> Parallel Processing Support

Decision tree induction has been widely used in extracting knowledge from feature based examples for classification and decision-making.

Fuzzy defined data are more accurate and it is able to analyze

the order in which different diagnostic tests should be performed. It can be used to minimize diagnostics costs and to guarantee a desired level of accuracy. The application of such estimations allows inducing minimum cost decision trees based on new optimality criteria. For this, a new type of fuzzy decision tree known as ordered tree was introduced. Ordered tree differs from unordered fuzzy decision tree in the way of testing attributes. In ordered trees the order of attribute tests is independent from the results of previous tests. This leads to decreasing of expenditures for test attributes. The problem of parallelizing the process of building a fuzzy decision tree can be treated stage by stage.

At the first step a fuzzification of the continuous data can be parallelized. The second improvement is possible through searching for maximum of the increment of information about the attribute at minimum cost. parallel processes can decrease the time for mining of association rules and programming model of the processes must be carefully selected to reflect their real configuration.

► **In Stock-Markets**

Time series data are of growing importance in many new database applications, such as data mining. A time series is a sequence of real numbers, each number representing a value at a time point. Fuzzy Decision Tree for dynamic stock exchange databases has been constructed by using weighted fuzzy production rules. In these rules, a weight parameter can be assigned to each proposition in the antecedent of a fuzzy production rule and certainty factor can be assigned to each rule. Certainty factors can be calculated by using some important variables in dynamic stock market.

Figure-3

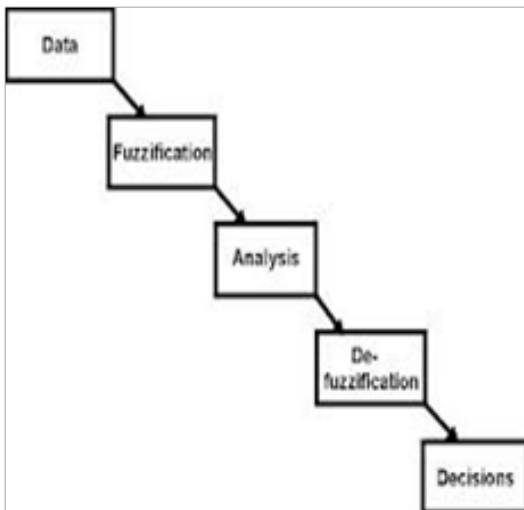
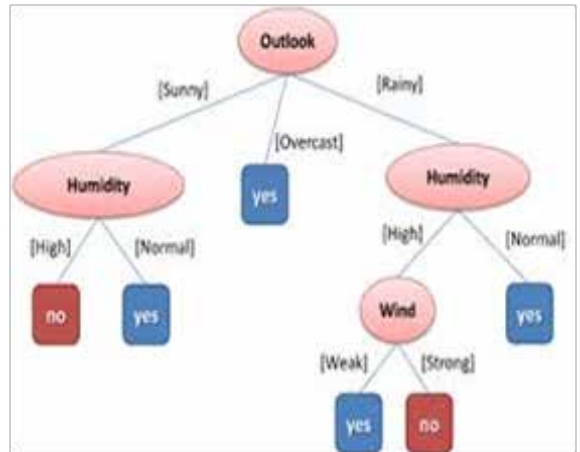


Figure-4



Conclusion

Thus Fuzzy Decision trees are one of the most popular choices for learning and reasoning. Graph theory has numerous applications in system analysis but in many cases some aspects of problems may be uncertain. In such cases it is natural to deal with uncertainty using fuzzy trees. For modeling complex systems using fuzzy graphs there are many advantages combine both fuzzy and graph applications.

REFERENCE

1.NarasimhaDeo, Graph Theory with Applications to science and Engineering | 2.George J. Kar Fuzzy sets Logic & Applications: phi Learning | 3. Jim-othy J.Ross, Fuzzy logic Engineering applications, McGrawhill | 4. George J.Klir, Fuzzy sets uncertainty & Information: phi Learning | 5. Fuzzy Graphs and Fuzzy Hypergraphs - In Depth Tutorials and Information | 6. Ferenc Peter Pach, Janos Abonyi, Supervised Clustering and Fuzzy Decision Tree, Induction for the Identification of Compact Classifiers | 7. Decision tree wikipedia free encyclopedia | 8. Robert Steiner: Fuzzy Logic in GIS |