Entropy and Algorithm of Obtaining Decision Trees in a Way Approximated to the Natural Intelligence

Olga Popova, Kuban State Technological University, Krasnodar, Russia
Boris Popov, Kuban State Technological University, Krasnodar, Russia
Vladimir Karandey, Kuban State Technological University, Krasnodar, Russia
Alexander Gerashchenko, Kuban State Technological University, Krasnodar, Russia

ABSTRACT
The classification of knowledge of a specified subject area is an actual task. The well-known methods of obtaining decision trees using entropy are not suitable for the classification of the subject area knowledge. So, a new algorithm of obtaining decision trees, whose way of obtaining is approximated to the natural intelligence, is suggested in the article. Here, the knowledge of a subject area is presented as a complex of answers to questions, which help to find the solution to a current task. The connection of entropy with the appearance of knowledge, the classification of previous knowledge, and the definitions used in decision trees are also analyzed in the article. The latter is necessary to compare the suggested algorithm approximated to the natural intelligence with the traditional method, using a small example. The article contains the analysis of solving a classification task for such a subject area as optimization methods.

KEYWORDS
Decision Tree, Entropy, Knowledge, Natural Intelligence, Production Rule

INTRODUCTION
Nowadays numerous classification tasks (Bahnsen, Aouada & Ottersten, 2015; Carlos & Abellán, 2014) are being solved in various areas of science. They have become actual due to the fast development of information technology. The progress in the methods of data collection, storage and processing has made it possible to collect huge masses of data. That is why the methods of automated analysis are necessary for such masses of data in order to process such a great amount of knowledge.

Currently, a number of popular problems can be solved with machine learning: classification, regression, clustering, anomalies detection and others. In classification problems, it is determined to which category the object can be referred according to its attributes. In regression problems, the attribute value of the object can be forecast on the basis of its other attributes. Clustering allows

DOI: 10.4018/IJCINI.2019070104

Copyright © 2019, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.
breaking down a set of objects into groups according to attributes of the objects in such a way that the objects were similar to each other within the groups but less similar beyond one group. In anomalies detection problems, search for objects that are “highly dissimilar” to all other ones in the sample or to a group of objects is performed. As for other problems, they are more specific. There are two types of machine learning algorithms – supervised (Deng, Wang, Li, Horng & Zhu, 2019; Cao, Qian, Wu & Wong, 2019) and unsupervised (Aissaoui, Madani, Oughdir & Alliou, 2019; Mak, Lee, & Park, 2019) ones. For unsupervised learning problems, they use samples consisting of objects that are described with a set of attributes. Meanwhile, supervised learning problems have an additional training sample for each object of which the target attribute is known. At present, two types of supervised learning problems are relevant: classification and regression ones. The simplest and the most popular method for solving these problems is the decision trees as they are used in daily life in the most diverse areas of human activity, and in ones being far away from machine learning at times. Visual directions of what to do in what situation can also be called a decision tree. The major advantage of decision trees consists in their being easy to interpret and similar to the model of human decision-making. This is why they have won immense popularity. So, classification method C4.5 using decision trees is considered as the first one in the list of top 10 data mining algorithms (Wu, Kumar, Quinlan, Ghosh, Yang, Motoda, McLachlan, Ng, Liu, Yu, Zhou, Steinbach, Hand, Steinberg, 2008).

Decision trees make it possible not only to classify all objects but also to obtain the tree structure with the optimal code (Abellán, Baker, Coolen, Crossman & Masegosa, 2014; Pei, Hu & Chen, 2016) having the minimal length. The demand for it grows with the increase in the number of classified objects, because the height of the tree increases as well. Here the resulting code is represented according to the Huffman tree type, by means of zeros and ones. For obtaining the optimal tree they use the information theory method suggesting the usage of the notions of information gain (Ferreira & Vasilyev, 2015; Abellán, 2013) and entropy (Nurpratami & Sitanggang, 2015; Kim, 2016). The obtained Shannon’s formula for the nodes of classification trees and the basic definitions taken from information theory provide a conventional way of building a decision tree. The way of choosing the attribute (Kamadi, Allam, Thummala & Rao, 2016; Ma, Destercke & Wang, 2016) and the rules for obtaining the nodes (Ferreira & Vasilyev, 2015; Abellán, López & de Oña, 2013) is always a creative process, which has the peculiarities conventional for this theory. Due to the fact that classification trees suggest adding the objects in the process of getting new experience, the problem of reconfiguring the tree is actual here. So it is necessary to develop the ways of teaching the decision support systems using such trees (Subburayalu, Jenhani & Slater, 2014; Cui, Shi & Wang, 2016). The resulting situation is as follows. There are no volume limits for describing the objects, their properties, facts, etc., which are to be classified and used for solving the complex scientific and applied tasks due to the modern advances in computer technology. The limits start to appear when the apparatus of analysis for these facts, objects, etc., is being arranged. These limits are the disadvantages of using the decision tree method, and they have to be urgently corrected. The authors suggest another way to eliminate the disadvantages of this method: they propose to change the conventional approaches to using it. The pre-existing approach to using the decision tree method did not make it possible to solve a number of tasks by the method of decision trees, which are especially “vulnerable” to using the “equivalent” structures of knowledge representation and the rules for obtaining the tree nodes (Popova, 2015; Popova, 2016; Popova, 2017; Popova, 2018a; Popova, 2018b). That is why the effect produced by automating such tasks should be greater than or at least equal to the efficiency of solving them manually. In this article, let us analyze one of such tasks and the way of solving it.

It is the task of the classification of knowledge for a given subject area. For solving such tasks, the authors suggest using a specific approach to obtaining decision trees and knowledge representation, which is approximated to the natural intelligence. Such an approach made it possible to obtain an automated system with the efficiency close to the work of the natural intelligence. For this purpose, it was necessary to analyze the efficient modern ways of structuring and working with information, which are used by the natural intelligence while obtaining the ideal and checked knowledge of a subject
Related Content

A Cognitive Computational Knowledge Representation Theory
Mehdi Najjar and André Mayers (2009). *Novel Approaches in Cognitive Informatics and Natural Intelligence* (pp. 247-264).
[www.igi-global.com/chapter/cognitive-computational-knowledge-representation-theory/27312?camid=4v1a](www.igi-global.com/chapter/cognitive-computational-knowledge-representation-theory/27312?camid=4v1a)

Fuzzy Neural Network Control for Robot Manipulator Directly Driven by Switched Reluctance Motor
[www.igi-global.com/article/fuzzy-neural-network-control-robot/60744?camid=4v1a](www.igi-global.com/article/fuzzy-neural-network-control-robot/60744?camid=4v1a)
Identity in the Real World
www.igi-global.com/chapter/identity-real-world/43693?camid=4v1a

Deductive Semantics of RTPA
www.igi-global.com/article/deductive-semantics-rtpa/1563?camid=4v1a