

LIMITED METHOD FOR THE CASE OF ALGORITHMIC CLASSIFICATION TREE

Povhan I. F. – PhD, Assistant Professor, Assistant Professor at the System Software Department, State Higher Education Institution Uzhhorod National University, Uzhhorod, Ukraine.

ABSTRACT

Context. The general problem of constructing the algorithmic recognition (classification) trees on the basis of a limited method in the artificial intelligence theory has been considered. The object of the present study is a concept of classification tree (a limited method-based algorithmic classification tree). The relevant methods, algorithms and schemes (a limited method) of constructing the algorithmic classification trees are the subject of this study.

Objective. The goal of this work is to develop a simple and efficient limited method of constructing the tree-like recognition and classification models on the basis of the algorithmic classification trees for training selection of a large-volume discrete information that is characterized by a structure of classification trees obtained from independent recognition algorithms assessed in accordance with their general efficiency calculation functional for a wide class of applied tasks.

Method. A limited method of constructing the algorithmic classification tree is suggested that constructs a tree-like structure for the preset initial training selection (an ACT model) consisting of a set of autonomous classification/recognition algorithms assessed at each ACT construction step (stage) in accordance with the initial classification. In other words, the limited method of constructing the algorithmic classification tree is suggested, and its idea is a step-by-step approximation of the arbitrary volume/structure selection by a set of independent classification/recognition algorithms. This method provides formation of a current algorithmic tree vertex (node, the generalized ACT attribute) with the selection of the most efficient (high-quality) autonomous classification algorithms from the initial set and construction completion of only those ACT structure paths, where the largest number of classification errors occurs. Such approach at constructing the resultant classification tree (the ACT model) allows the tree size and complexity (i.e. the total number of transitions, structure vertices and layers) to be reduced considerably, the quality of the next analysis (interpretability) and the possibility of decomposition to be increased as well as the ACT structures to be built given the limited hardware resources. The above limited method of constructing the algorithmic classification tree enables one to construct diverse tree-like recognition models with a preset accuracy for a wide class of the artificial intelligence theory tasks.

Results. The limited method of constructing the algorithmic classification tree developed and presented in this work has software realization and was investigated and compared to the logical classification tree methods (on the basis of elementary attribute set selection) and the algorithmic tree classification methods (first and second-type ones) when solving the task of real geological data recognition.

Conclusions. The experiments carried out in the present work have proved the performance capabilities of the software suggested and demonstrate the possibility of its promising utilization for the solution of a wide spectrum of applied recognition/classification problems. The outlook of further studies and approbations may be related to the creation of methods of other-type algorithmic classification trees that introduce a stopping criterion for the procedure of a tree model in accordance with the structure depth, optimization of its software realizations and to the experimental studies of this method for a wider circle of practical tasks.

KEYWORDS: algorithmic classification tree, image recognition, classification, classification algorithm, branching criterion, limited method.

ABBREVIATIONS

TS is a training selection;
ST is a test selection;
RS is a recognition system;
IR is a image recognition;
GA is a generalized attribute;
RF is a recognition function;
LCT is a logical classification tree;
ACT is a algorithmic classification tree;
GAT is a generalized attribute tree;
BAS is a branched attribute selection.

NOMENCLATURE

a_i is a fixed independent classification/recognition algorithms in the ACT scheme;
 m is a total number of training pairs (known classification objects) of the initial TS;
 Z is a limited method parameter that defines the maximal number of classification errors for a fixed path at the LCT construction completion;

M is a total number of independent classification algorithms a_i in a set;

T is a total number of the ST test sets;

G is a some initial signal (discrete object) manifold;

R is a partitioning into the classes (images) H_i defined at the initial manifold G ;

f_R is a recognition function RF defined at the manifold G ;

f_j is a fixed GAs constructed at the relevant ACT structure generation step;

x_i is a discrete objects (signals) of the initial TS;

H_i is a set of classes defined by the initial TS;

$(x_i, f_R(x_i))$ is a training pairs of the initial TS;

n is a total number of the problem attributes (the attribute space dimension);

k is a total number of classes of the manifold G ;

l is a value of the discrete object x belonging class;

$p_i(x)$ is a some predicate (a generalized attribute) for the i -th step of the LCT structure generation;

φ_i^j is an elementary x object attribute, here i being the attribute number in the set, j being the positioning layer number;

S is a number of the fixed path errors for the classification tree structure;

$r_1 r_2 r_3$ is a fixed attribute values defining some path in the LCT structure;

δ is a parameter that characterizes the training efficiency assessment with respect to the current problem/task;

$l(r_1, r_2, r_3)$ are quantity that characterizes the class belonging with respect to the fixed path in the classification tree structure;

$S_{r_1 r_2 r_3}$ is a number of all the pairs $(x_i, f_R(x_i))$ from the TS that, in fact, belong to the path $r_1 r_2 r_3$ and obey the relationship $f_R(x_i) \neq l(r_1 r_2 r_3)$;

G_{a_1, \dots, a_i} is a submanifold of the initial G approximated by a set of classification algorithms a_i ;

S_{a_1, \dots, a_n} is a number of the entrances into the TS of the initial pairs that satisfy condition $x_i \in G_{a_1, \dots, a_n}$;

$P_{pr}(TS)$ is a capacity (volume) of the initial TS (or its fixed part for a current ACT construction scheme step);

K_{Stop} is a branching stopping criterion at the ACT model construction;

En_{tr} is a classification errors and faults (the first-type ones) at the initial TS data array;

Et_{tr} is a classification errors and faults (the first-type ones) at the ST data array;

O_{Uz} is a total number of generalized attributes used in the classification tree model;

V_{All} is a represents the total number of all-type vertices in the ACT model structure;

N_{All} is a total number of different classification algorithms used in the classification tree model;

P_{All} is a total number of transitions between the vertices in the constructed classification tree model structure;

I_{Main} is a index of the initial TS data generalization by the classification tree;

Q_{Main} is a integral ACT/LCT model quality index;

Fr_{All} is a total number of vertices in the obtained ACT/LCT model with the resultant f_R values (RF, i.e. the classification tree leaves);

Er_{All} is a total number of errors in the ACT model at the initial test and training selection data arrays;

M_{All} is a total capacity (volume) of the test and training selection data arrays.

INTRODUCTION

The problems combined by the image recognition subject area are very diverse and take place in the contemporary world in any field of economy and social content of human activities resulting in the necessity to construct and study the mathematical model of the relevant systems. As of today, there is no universal approach to solve the above problems. Several quite general theories and approaches have been suggested allowing a large number of problem types/classes to be solved. However, their applied applications differ by rather high sensitivity to the problem or application field particularity. A lot of theoretical results have been obtained for a special cases and subtasks. It should be noted that the necessity of performing a huge amount of calculations and orientation to the powerful hardware tools remain a bottle-neck of successful real recognition systems. However a large number of applied tasks in various areas of nature studies, for instance, in geology, biology etc, where the problems of classification are being solved with the use of both software and hardware systems determines the intensity and relevance of such direction of investigations. Overcoming this problem for the tasks of artificial intelligence at the algorithmic and program construction of particular RS in a form of the classification tree models is a prerequisite of their high efficiency for each real task and, thus, will ensure a rapid development of different fields of science and technology [1–3]. In addition, more than four thousand recognition algorithms are available today being based on the various approaches and concepts and having certain limitations in their use (i.e. accuracy, speed, memory, versatility, reliability etc). Moreover, every algorithm is limited by a certain particularity of application tasks, and this is, obviously, the most vulnerable point of not such algorithms only, but also of the recognition systems based on the relevant concepts [4]. Thus, the logic classification trees (decision trees) are the object of this study. It is known that presentation of the large-volume training selections (discrete information) in a form of the logic tree structures has its essential advantages in economical description of the data and the efficient mechanisms of working with them [5–7].

The object of studies is the general concept of the decision tree, namely, the algorithmic classification tree, constructions (from the limited method point of view).

It is known that presentation of the large-volume discrete information arrays (training selections) in a form of the logic or algorithmic tree structures has its own essential advantages in economic description and efficient mechanisms (procedures) of working with them [8]. The effective covering of the training selection by a set of elementary attributes in case of the LCT structure or that by a fixed set of autonomous recognition/classification algorithms in case of the ACT construction generates a fixed tree-like data structure (a tree model) that ensures the training selection initial data compression and transformation. Let us emphasize that such approach provides considerable information system hardware resource optimization and savings [9]. The scope of using the decision

tree (the LCT/ACT structure) concept is now extremely vast, while the absolute majority of modern schemes of the classification tree construction methods are known from literature as partitioning and trapping [10–13] ones. It is known that the classification tree (LCT/ACT) structure is represented in a form of a sequence of branches and nodes. The tree branches contain some marks that define the target function (in case of the ACT, these are the independent classification algorithms, the GA sets), while the nodes contain the RF (independent class) values or the extended transition attributes. Choosing the branching criterion (vertex construction or selection), the branching stopping criterion (classification tree construction completion) and the classification tree branch (structural block) rejection criterion remain principal issues of the classification tree concept.

The subject of studies includes the methods, algorithms and construction schemes of classification algorithmic trees (the limited ACT structures).

The principal specific feature of the most available methods of training selection (discrete information array) processing in the recognition tasks at constructing classification rules and schemes is that they do not allow their complexity, accuracy and information capacity (parametric GA complexity) to be regulated in the course of model construction [4]. The principal specific feature of the algorithmic tree method is the possibility of the complex use of a number of known recognition algorithms (methods) in solving each particular task of recognition scheme construction. The basis of the ACT concept is a unique methodology, i.e. the optimal approximation of the training selection by a set of generalized attributes (autonomous algorithms) included in a certain scheme (operator) constructed during the training process [4, 5, 20]. The limited method of constructing the algorithmic classification tree models generates the tree-like schemes consisting of independent and autonomous classification algorithms presenting, to some extent, a new recognition algorithm (obviously, synthesized from known algorithms and methods).

The objective of the work is to elaborate the limited method of classification/recognition model construction on the basis of the ACT concept for the discrete information data arrays. Note that obtained classification system schemes are characterized by a tree-like structure (construction) and the availability of the autonomous classification algorithms (the GA sets) as the structural elements, whereas the above model construction (completing) goes along those tree structure paths, where the maximal number of classification errors takes place.

1 PROBLEM STATEMENT

Let the partition R into a finite number k of submanifolds (classes, images) be set at a certain manifold G of the objects $x: H_i (i=1, \dots, k)$, $G = \bigcup_{i=1}^k H_i$.

We shall call the corresponding manifolds H_1, \dots, H_k the images, while the manifold G elements – the patterns or the image H_1, \dots, H_k representatives. The objects (pat-

terns) x are defined by the certain attribute value sets $x_j (j=1, \dots, n)$. If $x \in H_i$, we shall consider that this object belongs to the image H_i . In general case, the images H_1, \dots, H_k can be defined by the probability distributions $p(H_1/x), \dots, p(H_k/x)$, here $p(H_i/x)$ being the probability (or, in the continuous case, the probability density) of $x(x \in G)$ belonging to the image H_i . Let the condition of task define a certain initial TS in a form of a sequence of training pairs:

$$(x_1, f_R(x_1)), \dots, (x_m, f_R(x_m)). \quad (1)$$

Note that, except for the initial TS, the ST (a set of the objects of known class belonging) is also defined as a certain part of the initial TS. Thus, according to the initial condition, TS is a combination (fixed sequence) of some sets (discrete objects), and each set is a combination of some attribute values and some function values (RF) at the above set. Then combination of attribute values is a certain pattern (discrete object), and the function value (RF) relates this pattern to the relevant image [5].

Therefore, in this work, we put a problem of constructing the ACT model in the conditions of limited hardware resources with the parameters p of the optimal structure $L: F(L(p, x_i), f_R(x_i)) \rightarrow opt$ with respect to the initial TS data.

2 REVIEW OF THE LITERATURE

Present study continues a cycle of works devoted to the problem of the tree-like discrete object recognition schemes (the LCT/ACT classification models) [7–9, 14, 20]. They deal with the principal issues of the classification tree structure construction, use and optimization. As known from Ref. [7], the resultant classification rule (scheme) constructed by an arbitrary BAS method or algorithm has a tree-like logic structure, and the logic tree consists of vertices (attributes) grouped in layers and obtained at a certain step (stage) of recognition tree construction [15]. So, for the ACT methods, an important problem that follows from Ref. [20] is a problem of synthesizing the recognition trees that would be represented, in fact, by the algorithm tree/graph (the ACT methods) or by the classification model tree. In contrary to the available methods, the main peculiarity of the tree-like recognition systems is that the importance of certain attributes (attribute or algorithm groups) is determined with respect to the function that defines the object partitioning into classes [16]. Thus, in Ref. [15], the principal issues concerning the decision tree generation and constructed model quality are analyzed for the case of low-informative attributes, and the ability of the classification tree structures to perform the one-dimensional branching (attribute selection) for the analysis of the importance/quality influence of certain variables (vertices) gives a chance to work with the various-type variables in the form of predicates, generalized attributes, while in case of the ACT – with the relevant autonomous classifi-

cation and recognition algorithms. The general classification tree concept is used actively in the intellectual data analysis (LGMB, XGBoost), where the final goal is the synthesis of the model (fixed scheme) that predicts the target variable value on the basis of the initial data set (the TS data arrays) at the system input [17].

The systems based on the CART methods (directed to the solution of classification and regressive analysis problems), as well as those on the basis of the C4.5 schemes and its modern modifications (for the recognition/classification problems) and the ID3 schemes are the approaches dominating in the decision tree concept. The ID3 scheme is based on using the limited entropy criterion, while the LCT structure is being constructed until for each resultant vertex (tree leaf) the objects of the same fixed class remain only, or the branching procedure in the tree under construction provides the initial entropy criterion reduction. The C4.5/C5.0 scheme is based on the well-known Gain-Ratio (normative entropy) criterion, and the limitation on the number of objects for the resultant vertex (the LCT structure leaf) is used as the procedure stopping criterion. Note that the pruning procedure in the LCT structure is carried out according to the Error-Based Pruning scheme based on the general assessment of the generalizing ability to make a decision on the classification tree construction branch/vertex elimination. The CART scheme uses in its work the Gini criteria, and the pruning procedure in the LCT structure is carried out according to the *Cost-Complexity Pruning* scheme, while, for the case of available attribute step over, the basic surrogate predicate scheme is used.

Note that the basic idea of the branched attribute (algorithm vertices) selection in the ACT structure could be determined as the optimal approximation of some initial TS by a set of the ranked classification algorithms (in case of the LCT – attributes). In this case the central issue, i.e. the problem of choosing the efficient branching criterion (vertex, attribute, discrete object attribute selection for the LCT and algorithm selection for the ACT) becomes the most important. These principal problems are analyzed in Refs. [20, 21], where the questions of quality assessment of particular discrete attributes, their sets and fixed connections are raised, and this allows the efficient mechanism of branching realization to be implemented.

It is known that the classification tree (LCT/ACT) model structures are characterized by a compactness, on the one side, and by a non-uniform layer filling (sparsity), on the other one, as compared to the regular tree constructions [18–24]. Here the issues of classification tree construction process using the methods of the branched attribute selection and those of choosing the logic tree synthesis stopping criterion remain relevant [25–34]. It should be noted that the classification tree concepts do not conflict with the possibility of using as the classification tree attributes (structural vertices) of not only certain attributes of their connecting objects (the generalized attribute idea was considered in Ref. [9]) but also the sets. If we do not consider the object attributes as branches but selects certain independent recognition algorithms, then

we get on the output a new structure, i.e. the ACT [20]. Present work will be devoted just to the ACT structures in view of the limited method.

3 MATERIALS AND METHODS

At this stage of the present work, we shall suggest the limited methods for the LCT and ACT (classification trees) structures that allow one to overcome some clear negative conceptual moments and limitations, which are inherent in such schemes, from the viewpoint of resource needs and the resultant complexity of constructed classification tree models [25].

Let us first draw some attention to the fact that the general scheme of the LCT model construction method (based on the step-by-step elementary attribute selection) described earlier in Ref. [21] has a principal shortcoming related to the fact that with the increase of the number of vertices (classification tree structure layers) in the LCT construction the number of elementary attributes φ_i^j (here i being the elementary attribute number in the set, j being the relevant attribute location layer number) increases significantly. Obviously, such the resulting LCT model complication (constructional complexity) affects negatively the hardware abilities of the classification system (memory, processor time) and the general ability to perceive and analyze the constructed model without external selection of classification rules in the tree structure. To overcome these principally negative moments of the classification tree method, we shall suggest the following LCT method modification.

Limited method of the LCT construction. Let us fix some positive number Z at the initial stage. Let us have a constructed LCT (the classification tree after the defined number of step of the LCT model construction) of a following general structure (see Fig. 1) that reflects some predicate (constructed generalized attribute) $p_1(x)$.

Note that in Ref. [9], when presenting the LCT construction method (on the basis of elementary attribute selection) at the test stage, a certain number S was calculated that appears in the following relation:

$$\frac{S}{m} \geq \delta. \quad (2)$$

Now we shall, for each unfinished path $r_1 r_2 r_3$ in the logic tree structure (Fig. 1), calculate, except for the number S , the other one – $S_{r_1 r_2 r_3}$ that expresses the number of all the pairs $(x_i, f_R(x_i))$ from TS that, in fact, belong to the path $r_1 r_2 r_3$ and obey the following relation:

$$f_R(x_i) \neq l(r_1 r_2 r_3). \quad (3)$$

Thus, $S_{r_1 r_2 r_3}$ is a number of all such errors due to some predicate $p_1(x)$ (generalized attribute) represented by this general-structure LCT (Fig. 1) at the fixed path $r_1 r_2 r_3$ in this tree construction.

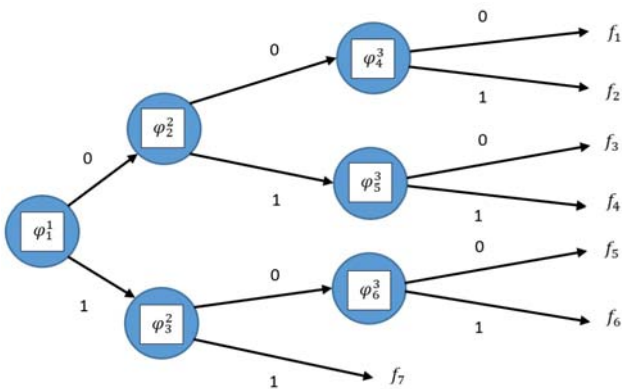


Figure 1 – Structure of the LCT constructed in accordance with the initial TS data on the basis of the elementary attribute selection

At the next stage, we choose the number Z of such paths $(r_1 r_2 r_3)_1, \dots, (r_1 r_2 r_3)_Z$, for which the number $S_{r_1 r_2 r_3}$ shall be as large as possible.

Example. Let $Z = 3$ be defined and let the following relation hold true:

$$S_{000} \geq S_{100} \geq S_{101} \geq S_{001} \geq S_{010} \geq S_{011}. \quad (4)$$

Then the paths 000, 100, 101 are chosen only. The next vertex (elementary attribute) $\varphi_{r_1 r_2 r_3}$ completion and selection shall be performed for the above paths only.

Let us fix that this modified scheme of the classification tree (the LCT on the basis of the elementary attribute selection) construction shall be called the limited method of the LCT construction.

Note that, according to this scheme, only the paths (of the general LCT structure) related to the maximal number of classification errors are being continued.

It should be noted finally that if one uses the above process in the end of the paths $r_1 r_2 r_3$ that do not belong to the selected Z paths, the $I(r_1 r_2 r_3)$ values are preserved. In this case, the process of the modified LCT construction method could be applied if the initial TS is not fixed, i.e. if at each step of the classification tree construction the own selection (a part of the TS) is realized.

Thus, the classification tree construction scheme suggested above introduces, actually, the accuracy regulation mechanism for the tree model constructed with the allowance made for the total number of classification errors at a given path (stage) of the logic tree general structure completion.

Obviously, this idea may work at the ACT structure level too, taking into account its certain specific features. Thus, in accordance with aforementioned, we shall suggest the following modification of the first-type ACT structure construction method presented in Ref. [20].

Limited method of the ACT construction. Let us define first a certain general-type TS (1) in a form of a sequence of training pairs $(x_i, f_R(x_i))$ with the capacity m , attribute space dimension n and a fixed set of different-type classification algorithms (a_1, \dots, a_M) . Note that

functioning of the classification tree models constructed is being checked at the ST data array with the capacity T (their belonging is also known).

It should be noted that here the initial TS data define some partition R into the classes (H_1, \dots, H_k) , while the corresponding algorithms a_i may not be associated with a single recognition concept but may realize various classification methods and algorithms (for instance, these could be the common geometrical algorithms that approximate the training selection by a relevant geometrical objects, i.e. estimate/potential function calculation algorithms etc). Note that one or more generalized attributes f_j (certain TS classification rules) that describe (approximate) a defined part of the initial training selection are the result of each of fixed (chosen from the algorithm library of some information system) autonomous classification/ recognition algorithms a_i at the relevant ACT generation step. Thus, for the case of known geometrical recognition algorithms [20], the geometrical objects that cover the TS in the attribute space of the dimension task space n will be the corresponding generalized attributes.

Certainly, in the real applied tasks, the cases are possible when the relevant classification algorithm a_i fails to construct the generalized attribute f_j due to a complicated location of classes H_k in the attribute space of this task or due to the certain conceptual and realization limitations of the classification algorithm itself. Then, by analogy with the LCT, we may deal with the case when constructed classification algorithms a_i (constructed generalized attributes f_j) incompletely approximate the initial TS, or such situation is provided by the ACT generation algorithm itself (for example, the presence of the initial limitation in the classification tree algorithm scheme related to generation of only one generalized attribute f_j at each stage of the ACT model construction).

Note that the initial TS objects that do not match the constructed scheme of selection approximation by a sequence of generalized attributes f_j (at the last stage of the ACT synthesis procedure) are related to the first-type classification failures (errors) En_{tr} . Similarly, for the ST data, the improperly classified discrete objects are also related to the first-type errors Et_{tr} .

Please note that the first-type ACT consists of layers, and each layer, in fact, corresponds to a certain step (stage) of the classification tree construction (the initial TS data approximation) [9, 20]. For each classification algorithm, at given approximation step, one may calculate its efficiency with respect the working data $(S/P_{pt}(TS^-))$, and this value must be equal to or exceed the preliminarily defined limitation δ . Certainly, in some ACT scheme realizations, this value δ may be used as

the stopping criterion K_{Stop} for the classification tree structure branching procedure.

Note that here S is a total number of classification errors for the fixed algorithm at a certain step (the ACT generation stage), while $P_{pr}(TS^-)$ is a capacity (volume) of the initial TS submanifold fed to the given algorithm input at the relevant layer (level or stage) of the classification tree under construction.

Then, by analogy with the limited LCT method, it looks expedient to calculate for each ACT structure layer (the classification tree construction stage) the values S_{a_1, \dots, a_i} that characterize the number of all the pairs $(x_i, f_R(x_i))$ of the initial TS array that could not be approximated by a sequence of fixed classification algorithms a_1, \dots, a_i . Thus, S_{a_1, \dots, a_i} is a number of all the classification errors made by a certain sequence GA (constructed at the corresponding ACT structure levels) represented by a given ACT of the general structure (Fig. 2) for a fixed recognition/classification algorithm set a_1, \dots, a_i .

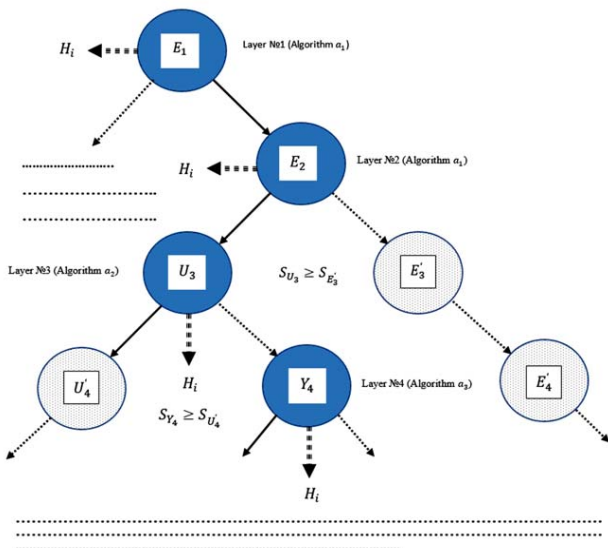


Figure 2 – Classification tree structure fragment constructed using the limited ACT method

Then, according to this ACT structure construction scheme, the only path(s) should be chosen in the tree construction (dependent of the classification tree types that may differ), for which the value S_{a_1, \dots, a_i} will be as maximal as possible. That is, the path in the ACT structure with the largest number of errors is being completed. Note that the next completion and selection of vertices (classification algorithms – the GA parameters) a_1, \dots, a_i is performed for these paths in the ACT structure only.

Note that in such approach to the construction of the limited LCT/ACT methods, after constructing the classification tree model its actual completion (final training) is possible and this provides a direct possibility to influence the accuracy of the classification system model constructed.

It should be noted that, according to the limited ACT scheme, in the process of the classification tree construction, only those paths (of the general ACT structure) are being continued (chosen for the GA completion) that correspond to the occurrence of the largest number of classification errors. It is obvious that the use of the above process allows one to return at any moment to the rejected paths in order to complete the classification tree structures (paths).

Let us emphasize that the classification tree construction scheme suggested above (the limited ACT method) allows the constructed tree model accuracy (efficiency) to be regulated taking into account the total number of any-type classification errors at a given path (stage) of the algorithmic tree general structure construction. The principal possibility of the ACT model construction with a preset accuracy with respect to the initial data array is important. Such possibility is achieved by limiting/ restricting the number of steps of the ACT generation procedure, by using a system of limitations of information capacity, the generalization number and parameters (the TS domain under approximation) or the generalized attribute set constructed at the relevant stages of the resultant classification tree construction.

4 EXPERIMENTS

The limited ACT structure construction scheme suggested in this study allows one to regulate quite flexibly the complexity of the model of the classification tree under construction or to construct the recognition model with the preset accuracy in accordance with the current task conditions. The task of choosing the classification tree model (among a fixed set of constructed LCT/ACT structures) for a particular problem is determined by a manifold of parameters that have a determinative importance with respect to the current applied task (the TS/ST data sets). It is obvious that, in order to compare and select a particular classification tree model from the fixed set, one has to distinguish its most significant characteristics (i.e. attribute space dimensionality, number of vertices, number of the tree construction transitions etc) and to determine their error with respect to the input data array.

Analyzing the quality criteria for the information models obtained that depend on the model inaccuracy, the initial TS/ST data array capacity (the number of training pairs and the task attribute space dimensionality), the number of the model structural parameters and so on, is the principally important issue at this stage of study. It is obvious that the model errors at the TS and ST data arrays and for each class (the initial TS part, submanifold) defined by the current applied task initial condition are the critically important parameters of the ACT model constructed.

Note that the basic index of the initial TS data generalization by the classification tree/model is one of the most important indices that characterize the basic properties of the ACT models. It is calculated as follows:

$$I_{Main} = \frac{m \cdot O_{Uz}}{V_{All} + N_{All} + 2P_{All}}. \quad (5)$$

This classification tree model (the ACT structure) generalization index reflects its basic parameters/ characteristics of classification trees and may be used as an optimality criterion in the procedure of assessing the arbitrary tree-like recognition scheme, for example, in the case of the methods of constructing and selecting the random classification trees from Ref. [9] (with the allowance made for their structural parameters). In any arbitrary applied task, it is essential to maximize the parameter I_{Main} (i.e. the ACT model generalization index). This allows one to achieve the most optimal classification tree structure and obtain actually the maximal compression of the initial TS data (i.e. to present the initial data array as the minimal by its structural complexity tree) [7]. It should also be emphasized that the issue of reducing the tree structure complexity (namely, the number of nodes, vertices, attributes, algorithms in the classification tree structure, total number of transitions in the model structure), the information system general memory and processor time consumption parameters remains a principal moment at the LCT/ACT (classification model) structure construction. So, the total integral quality index presented below is an important quality index of the model constructed in a form of the classification tree with the inclusion of the ACT model structure parameters:

$$Q_{Main} = \frac{Fr_{All}}{O_{Uz} \cdot \sum_i p_i} \cdot e^{-\frac{Er_{All}}{M_{All}}} \quad (6)$$

Note that the parameter Er_{All} is here the total number of errors in the ACT model at the initial test and training selection data arrays – $Er_{All} = En_{tr} + Et_{tr}$, while M_{All} is the total capacity of the above two arrays – $M_{All} = m + T$. The parameter Fr_{All} characterizes the total number of vertices in the obtained ACT model with the resultant values f_R (RF, i.e. the classification tree leaves), while the parameter O_{Uz} represents the total number of all the generalized attributes in the ACT model structure. A set of parameters p_i defines the most important classification tree characteristics (in accordance with the LCT/ACT structures) under assessment (for instance, the number of elementary or generalized attributes used in the classification tree model, the number of transitions between the classification tree vertices, layers etc).

Note that this integral ACT model quality index will take the values from zero to unit. The less is this index, the worse is the quality of the classification tree constructed, and, vice versa, the larger is the index, the best will be the final model. Thus, the suggested integral assessment of the classification tree (the ACT structure) reflects the basic classification tree parameters/characteristics and may be used as the optimality criterion in the procedure of assessment of the arbitrary tree-like recognition scheme (according to the own model parameters).

Thus, at the Uzhhorod National University, the software complex Orion III has been developed being used to generate the autonomous recognition/classification systems. This system algorithmic library has 13 recognition algorithms/schemes including the above algorithmic realization of the ACT construction using the limited method suggested above.

The basic task used to check operation of the limited ACT construction method was reduced to that of classifying the array of geological data (the oil-bearing bed partitioning problem). A set of 22 basic elementary attributes was used to recognize objects, whereas in the TS information was presented about the objects from two classes. At the examination stage, the classification system (the ACT model) was constructed being intended to provide an efficient recognition of unknown classification objects with respect to these two classes.

Information concerning the two classes of objects is presented in the TS. At the examination stage, the classification system constructed has to provide an efficient recognition of unknown classification methods with respect to the above two classes.

At the initial stage of the software system operation, the training selection was automatically checked for correctness (searching and eliminating the similar objects of different belonging – the first-kind errors).

5 RESULTS

The initial TS presents information about the partition R into two classes. At the examination stage (on the ST basis), the classification scheme constructed must ensure the efficient recognition (classification) of unknown-classification objects with respect to these classes. Note that the H_1 -class training pairs (the oil-bearing beds) dominated in the training information array in the proportion $\approx (1.5/1)$, while the TS array included 1250 objects (sets of known classification), and the efficiency of the constructed recognition system was estimated at the test volume selection of 240 objects. It should be noted that the ST array had a form of a separated part of the initial TS (it consisted of discrete objects of known classification). Usually such test selection volume is not sufficient for a comprehensive analysis of the constructed classification tree model quality, but in relation with the limited character of the TS itself even such ST allows the main parameters of synthesized LCT/ACT structures to be assessed and analyzed. The test and training selection array data were obtained on the basis of the geological survey on the territory of Transcarpathian province during the period from 2001 to 2011. The fragment of the main results of the above experiments, the comparative tests of the ACT (the classification tree structure) model construction methods at the data array for this applied task is presented in Table 1. We shall emphasize that constructed ACT models (structures) provided necessary accuracy and efficiency level defined by the applied task condition, desired speed and system operating memory consumption, but demonstrated different structural complexity of constructed classification trees (the LCT construction

complexity parameters) and a set of generalized attributes (in the comparative cases of the algorithmic classification tree models, i.e. the ACT structures [20]).

6 DISCUSSION

Note that the general assessment of the ACT model quality made in the present study (certainly, it may be adapted for the case of the LCT structures) fixes the most important characteristics/parameters of the constructed classification trees and can be used as the optimality criterion in the procedure of the ACT construction and final model selection (ranking) from the ACT model manifold. It should be noted that the ACT structure, like the limited algorithmic tree method, only handles the constructed sets of generalized attributes (classification algorithms) and may not be interested in the fact by what algorithm/method (scheme, rule) they have been obtained. Each of the schemes constructed according to the algorithmic tree method will be a general recognition scheme (the ACT model) to be used for practical purposes (i.e. processing large experimental data arrays in a form of discrete sets). It is important that the classification scheme obtained will, to some extent, be a new recognition algorithm (synthesized, obviously, from the known algorithms and methods). The ACT structure (new classification scheme) obtained is characterized by a high flexibility with respect to a certain application and by a relatively compact structure of the model itself, however, it requires relatively high hardware expenditures to preserve generalized attributes (or their sets) and the initial assessment of the classification algorithm quality according to the TS data. The ACT models, as compared to the LCT structures, have high classification rule speed, comparable hardware expenditures for the tree structure preservation and operation, as well as high classification quality.

CONCLUSIONS

In this work, a problem of constructing the limited method of synthesizing the algorithmic classification tree models on the basis of the TS approximation by a set of independent classification algorithms in conditions of limitations imposed on the paths of the ACT structure completion has been solved.

The **scientific novelty** of the obtained results is related to the fact that the limited method of constructing the ACT structures has been suggested for the first time on the basis of the autonomous recognition/classification algorithm assessment and ranking for generating the classification tree structure (the ACT model) with the limitation set imposed on the ACT (the classification tree model) construction completion directions. In this case at each step of the classification tree branching a certain part of the TS (or its submanifold) is being approximated. Note that the branching criterion for the ACT structures (in the limited ACT construction method) may be used not only to assess the quality of some classification algorithms, but also to calculate the efficiency of bound algorithm sets that will allow more optimal structure of the ACT synthesized in accordance with the initial TS data to be achieved. In this work, a set of general indices (parameters) has been suggested enabling the general characteristics of the ACT model to be presented effectively. It also could be used to select the most optimal ACT from the set of classification trees constructed in accordance with the random classification trees.

The **practical value** of the results obtained is that the suggested limited method of constructing the ACT models (the LCT/ACT structures) was realized in the algorithm library of the universal software system ORION III to solve various practical problems of classifying (recognizing) the different-type discrete object arrays.

Note that the practical testing has proven the efficiency and performance of the developed software and suggested ACT models allowing one to formulate recommendations on the use of the above approach (the limited LCT/ACT model method) and its software realization for a wide spectrum of applied discrete object classification/recognition tasks.

Looking ahead, the further studies may be directed towards the development of algorithmic classification tree (the ACT structure boosting) methods of optimizing the software realizations of the suggested limited ACT construction method as well as its practical approbation at the family of real classification/recognition tasks.

Table 1 – Comparative table of classification tree (LCT/ACT) models/methods

No.	Classification tree (LCT/ACT) method (scheme)	Integral classification tree model quality index, Q_{Main}	Total number of model errors at TS and ST, Er_{All}
1	Full LCT method on the basis of elementary attribute (attribute set) selection	0.004789	2
2	LCT model with a single assessment of attribute (attribute set) importance	0.002263	3
3	Limited LCT structure construction method ($Z=5$)	0.003168	4
4	Limited LCT structure construction method LCT ($Z=9$)	0.003029	3
5	Algorithmic tree (type I) method	0.005234	0
6	Algorithmic tree (type II) method	0.002941	0
7	ACT method (type I) on the basis of a hypersphere algorithm as the GA	0.005445	0
8	ACT method (type I) on the basis of a hypercube algorithm as the GA	0.005139	1
9	Limited ACT model construction method ($Z=6$)	0.003018	2
10	Limited ACT model construction method ($Z=10$)	0.003176	1

REFERENCES

1. Srikant R., Agrawal R. Mining generalized association rules, *Future Generation Computer Systems*, 1997, Vol. 13, No. 2, pp. 161–180.
2. Hastie T., Tibshirani R., Friedman J. The Elements of Statistical Learning. Stanford, 2008, 768 p.
3. Quinlan J.R. Induction of Decision Trees. Machine Learning, 1986, No. 1, pp. 81–106.
4. Vasilenko Y. A., Vasilenko E. Y., Kuhayivsky A. I., Papp I. O. Construction and optimization of recognizing systems, *Scientific and technical journal "Information technologies and systems"*, 1999, No. 1, pp. 122–125.
5. Povhan I. Designing of recognition system of discrete objects, *2016 IEEE First International Conference on Data Stream Mining & Processing (DSMP), Lviv, 2016. Ukraine, Lviv, 2016*, pp. 226–231.
6. Mitchell T. Machine learning. New York, McGrawHill, 1997, 432 p.
7. Povhan I. General scheme for constructing the most complex logical tree of classification in pattern recognition discrete objects, *Collection of proceedings «Electronics and information technology»*, 2019, Vol. 11, pp. 73–80.
8. Breiman L. L., Friedman J. H., Olshen R. A. et al. Classification and regression trees. Boca Raton, Chapman and Hall/CRC, 1984, 368 p.
9. Vasilenko Y. A., Vashuk F. G., Povkhan I. F. Automating the construction of classification systems based on agent – schemes, *Mathematical modeling, optimization and information technologies : International Joint Conference MDIF-2012, Kishineu, Moldova, 2012*. Kishineu, 2012, pp. 444–446.
10. Vtoghoff P.E. Incremental Induction of Decision Trees, *Machine Learning*, 1989, No. 4, pp. 161–186.
11. Amit Y., Geman D., Wilder K. Joint induction of shape features and tree classifiers, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997, Vol. 19, No. 11, pp. 1300–1305.
12. Dietterich T. G., Kong E. B. Machine learning bias, statistical bias and statistical variance of decision tree algorithms [Electronic resource]. Corvallis, Oregon State University, 1995, 14 p. Access mode : <http://www.cems.uwe.ac.uk/~irjohnso/coursenotes/uqc832/trbias.pdf>
13. Mingers J. An empirical comparison of pruning methods for decision tree induction, *Machine learning*, 1989, Vol. 4, No. 2, pp. 227–243.
14. Povhan I. Question of the optimality criterion of a regular logical tree based on the concept of similarity, *Collection of proceedings «Electronics and information technology»*, 2020, Vol. 13, pp. 12–16.
15. Subbotin S.A. Construction of decision trees for the case of low-information features, *Radio Electronics, Computer Science, Control*, 2019, No. 1, pp. 121–130.
16. Lupei M., Mitsa A., Repariuk V., Sharkan V. Identification of authorship of Ukrainian-language texts of journalistic style using neural networks, *Eastern-European Journal of Enterprise Technologies*, 2020, Vol. 1 (2 (103)), pp. 30–36. DOI: <https://doi.org/10.15587/1729-4061.2020.195041>
17. Bodyanskiy Y., Vynokurova O., Setlak G. and Pliss I. Hybrid neuro-neo-fuzzy system and its adaptive learning algorithm, *Computer Sciences and Information Technologies (CSIT) : Xth Scien. and Tech. Conf.*, Lviv, 2015. Lviv, 2015, pp. 111–114.
18. Karimi K., Hamilton H. J. Generation and Interpretation of Temporal Decision Rules, *International Journal of Computer Information Systems and Industrial Management Applications*, 2011, Vol. 3, pp. 314–323.
19. Kotsiantis S.B. Supervised Machine Learning: A Review of Classification Techniques, *Informatica*, 2007, No. 31, pp. 249–268.
20. Povkhan I.F. Features of synthesis of generalized features in the construction of recognition systems using the logical tree method, *Information technologies and computer modeling ITKM-2019 : materials of the international scientific and practical conference, Ivano-Frankivsk, May 20–25, 2019, Ivano-Frankivsk, 2019*, pp. 169–174.
21. Vasilenko Y. A., Vashuk F. G., Povkhan I. F. The importance of discrete signs, *XX International Conference Promising ways and directions of improving the educational system, Uzhgorod, November 16–19, 2010*. Uzhgorod, 2010, Vol. 21, No. 1, pp. 217–222.
22. Deng H., Runger G., Tuv E. Bias of importance measures for multi-valued attributes and solutions, *Proceedings of the 21st International Conference on Artificial Neural Networks (ICANN), Espoo, Finland, Jun 14–Jun 17, 2011*. Espoo, 2011, pp. 293–300.
23. Kamiński B., Jakubczyk M., Szufel P. A framework for sensitivity analysis of decision trees, *Central European Journal of Operations Research*, 2017, Vol. 26 (1), pp. 135–159.
24. Dietterich T. G. An experimental comparison of three methods for constructing ensembles of decision trees: bagging, boosting, and randomization, *Machine learning*, 2000, Vol. 40, No. 2, pp. 139–157.
25. Povhan I. Generation of elementary signs in the general scheme of the recognition system based on the logical tree, *Collection of proceedings «Electronics and information technology»*, 2019. – Vol. 12. – P. 20–29.
26. Subbotin S., Oliinyk A. eds.: Szewczyk R., Kaliczyńska M. The dimensionality reduction methods based on computational intelligence in problems of object classification and diagnosis, *Recent Advances in Systems, Control and Information Technology*. Cham, Springer, 2017, pp. 11–19. (Advances in Intelligent Systems and Computing, vol. 543).
27. Subbotin S. A. Methods and characteristics of locality-preserving transformations in the problems of computational intelligence, *Radio Electronics, Computer Science, Control*, 2014, No. 1, pp. 120–128.
28. Koskimaki H., Juutilainen I., Laurinen P., Roning J. Two-level clustering approach to training data instance selection: a case study for the steel industry, *Neural Networks : International Joint Conference (IJCNN-2008), Hong Kong, 1–8 June 2008 : proceedings*. Los Alamitos, IEEE, 2008, pp. 3044–3049. DOI: 10.1109/ijcnn.2008.4634228
29. Subbotin S. The neuro-fuzzy network synthesis and simplification on precedents in problems of diagnosis and pattern recognition, *Optical Memory and Neural Networks (Information Optics)*, 2013, Vol. 22, No. 2, pp. 97–103. DOI: 10.3103/s1060992x13020082
30. Subbotin S. A. Methods of sampling based on exhaustive and evolutionary search, *Automatic Control and Computer Sciences*, 2013, Vol. 47, No. 3, pp. 113–121. DOI: 10.3103/s0146411613030073
31. De Mántaras R. L. A distance-based attribute selection measure for decision tree induction, *Machine learning*, 1991, Vol. 6, No. 1, pp. 81–92.
32. Alpaydin E. Introduction to Machine Learning. London, The MIT Press, 2010, 400 p.

33. Painsky A., Rosset S. Cross-validated variable selection in tree-based methods improves predictive performance, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, Vol. 39, No. 11, pp. 2142–2153. DOI:10.1109/tpami.2016.2636831
34. Miyakawa M. Criteria for selecting a variable in the construction of efficient decision trees, *IEEE Transactions on Computers*, 1989, Vol. 38, № 1, pp. 130–141. Received 22.06.2020. Accepted 14.09.2020.

УДК 001.891:65.011.56

ОБМЕЖЕНИЙ МЕТОД ДЛЯ ВПАДКУ АЛГОРИТМИЧНОГО ДЕРЕВА КЛАСИФІКАЦІЇ

Повхан І. Ф. – канд. техн. наук, доцент, доцент кафедри програмного забезпечення систем ДВНЗ Ужгородський національний університет, м. Ужгород, Україна.

АНОТАЦІЯ

Актуальність. Розглянута загальна задача побудови алгоритмічних дерев розпізнавання (класифікації) на основі обмеженого методу в теорії штучного інтелекту. Об'єктом даного дослідження є концепція дерева класифікації (алгоритмічного дерева класифікації на базі обмеженого методу). Предметом дослідження є актуальні методи, алгоритми та схеми (обмежений метод) побудови алгоритмічних дерев класифікації.

Мета. Метою даної роботи є створення простого та ефективного обмеженого методу побудови деревоподібних моделей розпізнавання та класифікації на основі алгоритмічних дерев класифікації для навчальних вибірок дискретної інформації великого об'єму – який характеризується структурою отриманих дерев класифікації з незалежних алгоритмів розпізнавання оцінених на основі функціоналу розрахунку їх загальної ефективності для широкого класу прикладних задач.

Метод. Пропонується обмежений метод побудови алгоритмічних дерев класифікації, який для заданої початкової навчальної вибірки довільного розміру будує деревоподібну структуру (модель АДК), яка складається з набору автономних алгоритмів класифікації та розпізнавання оцінених на кожному кроці (етапі) побудови АДК за даною початковою вибіркою. Тобто пропонується обмежений метод побудови алгоритмічного дерева класифікації основна ідея якого полягає в по кроковій апроксимації початкової вибірки довільного об'єму та структури набором незалежних алгоритмів класифікації та розпізнавання. Даний метод при формуванні поточної вершини алгоритмічного дерева (вузла, узагальненої ознаки АДК) забезпечує виділення найбільш ефективних (якісних) автономних алгоритмів класифікації з початкового набору та побудову лише тих шляхів в структурі АДК де відбувається найбільша кількість помилок класифікації. Такий підхід при побудові результуючого дерева класифікації (моделі АДК) дозволяє значно скоротити розмір та складність дерева (загальну кількість переходів, вершин та ярусів структури) підвищити якість його наступного аналізу (інтерпретабельність), можливість декомпозиції, та будувати структури АДК в умовах обмежених апаратних ресурсів. Запропонований обмежений метод побудови алгоритмічного дерева класифікації дозволяє будувати різноманітні деревоподібні моделі розпізнавання з наперед заданою точністю для широкого класу задач теорії штучного інтелекту.

Результати. Розроблений та представлений в даній роботі обмежений метод алгоритмічного дерева класифікації отримав програмну реалізацію та був досліджений і порівняний з методами логічних дерев класифікації (на основі селекції набору елементарних ознак), методами алгоритмічного дерева класифікації (першого та другого типу) при розв'язку задачі розпізнавання реальних даних геологічного типу.

Висновки. Проведені в даній роботі експерименти підтвердили працездатність запропонованого математичного забезпечення та показують можливість його перспективного використання для розв'язку широкого спектру практичних задач розпізнавання та класифікації. Перспективи подальших досліджень та апробацій можуть полягати в створенні методів алгоритмічного дерева класифікації інших типів, які полягають в веденні критерію зупинки процедури побудови моделі дерева за глибиною структури, оптимізації його програмних реалізацій, а також експериментальних дослідженнях даного методу на більш широке коло практичних задач.

КЛЮЧОВІ СЛОВА: алгоритмічне дерево класифікації, розпізнавання образів, класифікація, алгоритм класифікації, критерій розгалуження, обмежений метод.

УДК 001.891:65.011.56

ОГРАНИЧЕННЫЙ МЕТОД ДЛЯ СЛУЧАЯ АЛГОРИТМИЧЕСКОГО ДЕРЕВА КЛАССИФИКАЦИИ

Повхан И. Ф. – канд. техн. наук, доцент, доцент кафедры программного обеспечения систем ГВУЗ Ужгородский национальный университет, г. Ужгород, Украина.

АННОТАЦИЯ

Актуальность. Рассмотрена общая задача построения алгоритмических деревьев распознавания (классификации) на основе ограниченного метода в теории искусственного интеллекта. Объектом данного исследования является концепция дерева классификации (алгоритмического дерева классификации на базе ограниченного метода). Предметом исследования являются актуальные методы, алгоритмы и схемы (ограниченный метод) построения алгоритмических деревьев классификации.

Цель. Целью данной работы является создание простого и эффективного ограниченного метода построения древовидных моделей распознавания и классификации на основе алгоритма деревьев классификации для учебных выборок дискретной информации большого объема – характеризуется структурой полученных деревьев классификации из независимых алгоритмов распознавания оцененных на основе функционала расчета их общей эффективности для широкого класса прикладных задач.

Метод. Предлагается ограниченный метод построения алгоритмических деревьев классификации, который для заданной начальной обучающей выборки произвольного размера строит древовидную структуру (модель АДК), которая состоит из набора автономных алгоритмов классификации и распознавания, оцененных на каждом шаге (этапе) построения АДК по данной начальной выборке. То есть предлагается ограниченный метод построения алгоритмического дерева классификации основной идея которого заключается в по шаговой аппроксимации начальной выборки произвольного объема и структуры набором независимых алгоритмов классификации и распознавания. Данный метод при формировании текущей вершины алгоритмического дерева (узла, обобщенной признаки АДК) обеспечивает выделение наиболее эффективных (качественных) автономных алгоритмов классификации с начального набора и достройку только тех путей в структуре АДК где происходит наибольшее количество ошибок классификации. Такой подход при построении результирующего дерева классификации (модели АДК) позволяет значительно сократить размер и сложность дерева (общее количество переходов, вершин и ярусов структуры) во-визжит качество его последующего анализа (интерпретируемость), возможность декомпозиции, и строить структуры АДК в условиях ограниченных аппаратных ресурсов. Предложенный ограниченный метод построения алгоритмического дерева классификации позволяет строить разнотипные древовидные модели распознавания с наперед заданной точностью для широкого класса задач теории искусственного интеллекта.

Результаты. Разработан и представлен в данной работе ограничен метод алгоритмического дерева классификации получил программную реализацию и был исследован и сравнен с методами логических деревьев классификации (на основе селекции набора элементарных признаков), методами алгоритмического дерева классификации (первого и второго типа) при решении задачи распознавания реальных данных геологического типа.

Выводы. Проведенные в данной работе эксперименты подтвердили работоспособность предложенного математического обеспечения и показывают возможность его перспективного использования для решения широкого спектра практических задач распознавания и классификации. Перспективы дальнейших исследований и апробаций могут заключаться в создании методов алгоритмического дерева классификации других типов, которые заключаются в ведении критерия останковки процедуры построения модели дерева по глубине структуры, оптимизации его программных реализаций, а также экспериментальных исследованиях данного метода на более широкий круг практических задач.

КЛЮЧЕВЫЕ СЛОВА: алгоритмическое дерево классификации, распознавание образов, классификация, алгоритм классификации, критерий разветвления, ограниченный метод.

REFERENCES

1. Srikant R. Mining generalized association rules / R. Srikant, R. Agrawal // *Future Generation Computer Systems*. – 1997. – Vol. 13, №2. – P. 161–180.
2. Hastie T. *The Elements of Statistical Learning* / T. Hastie, R. Tibshirani, J. Friedman. – Stanford, 2008. – 768 p.
3. Quinlan J. R. *Induction of Decision Trees* / J. R. Quinlan // *Machine Learning*. – 1986. – № 1. – P. 81–106.
4. Construction and optimization of recognizing systems / [Vasilenko Y. A., Vasilenko E. Y., Kuhayivsky A. I., Papp I. O.] // *Scientific and technical journal “Information technologies and systems”*. – 1999. – № 1. – P. 122–125.
5. Povhan I. Designing of recognition system of discrete objects / I. Povhan // *2016 IEEE First International Conference on Data Stream Mining & Processing (DSMP)*, 2016, Lviv, Ukraine. – Lviv, 2016. – P. 226–231.
6. Mitchell T. *Machine learning* / T. Mitchell. – New York : McGraw-Hill, 1997. – 432 p.
7. Povhan I. General scheme for constructing the most complex logical tree of classification in pattern recognition discrete objects / I. Povhan // *Collection of proceedings «Electronics and information technology»*. – 2019. – Vol. 11. – P. 73–80.
8. *Classification and regression trees* / [L. L. Breiman, J. H. Friedman, R. A. Olshen, C. J. Stone]. – Boca Raton : Chapman and Hall/CRC, 1984. – 368 p.
9. Vasilenko Y. A. Automating the construction of classification systems based on agent – schemes / [Y. A. Vasilenko, F. G. Vashuk, I. F. Povkhan] // *Mathematical modeling, optimization and information technologies : International Joint Conference MDIF-2012, Kishineu, Moldova, 2012*. – Kishineu, 2012. – P. 444–446.
10. Vtoghoff P. E. Incremental Induction of Decision Trees / P. E. Vtoghoff // *Machine Learning*. – 1989. – № 4. – P. 161–186.
11. Amit Y. Joint induction of shape features and tree classifiers / Y. Amit, D. Geman, K. Wilder // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. – 1997. – Vol. 19, № 11. – P. 1300–1305.
12. Dietterich T. G. Machine learning bias, statistical bias, and statistical variance of decision tree algorithms [Electronic resource] / T. G. Dietterich, E. B. Kong. – Corvallis : Oregon State University, 1995. – 14 p. – Access mode : <http://www.cems.uwe.ac.uk/~irjohnso/coursenotes/uqc832/trbias.pdf>
13. Mingers J. An empirical comparison of pruning methods for decision tree induction / J. Mingers // *Machine learning*. – 1989. – Vol. 4, No. 2. – P. 227–243.
14. Povhan I. Question of the optimality criterion of a regular logical tree based on the concept of similarity / I. Povhan // *Collection of proceedings «Electronics and information technology»*. – 2020. – Vol. 13. – P. 12–16.
15. Subbotin S. A. Construction of decision trees for the case of low-information features / S. A. Subbotin // *Radio Electronics, Computer Science, Control*. – 2019. – № 1. – P. 121–130.
16. Identification of authorship of Ukrainian-language texts of journalistic style using neural networks / [M. Lupei, A. Mitsa, V. Repariuk, V. Sharkan] // *Eastern-European Journal of Enterprise Technologies*. – 2020. – Vol. 1 (2 (103)). – P. 30–36. DOI: <https://doi.org/10.15587/1729-4061.2020.195041>
17. Hybrid neuro-neo-fuzzy system and its adaptive learning algorithm / [Y. Bodyanskiy, O. Vynokurova, G. Setlak and I. Pliss] // *Computer Sciences and Information Technologies (CSIT) : Xth Scien. and Tech. Conf., Lviv, 2015*. – Lviv, 2015. – P. 111–114.
18. Karimi K. Generation and Interpretation of Temporal Decision Rules / K. Karimi, H. J. Hamilton // *International Journal of Computer Information Systems and Industrial Management Applications*. – 2011. – Vol. 3. – P. 314–323.
19. Kotsiantis S. B. Supervised Machine Learning: A Review of Classification Techniques / S. B. Kotsiantis // *Informatica*. – 2007. – № 31. – P. 249–268.

20. Povkhan I.F. Features of synthesis of generalized features in the construction of recognition systems using the logical tree method / I. F. Povkhan // *Information technologies and computer modeling ITKM-2019 : materials of the international scientific and practical conference, Ivano-Frankivsk, May 20–25, 2019. – Ivano-Frankivsk, 2019. – P. 169–174.*
21. Vasilenko Y. A. The importance of discrete signs / Y. A. Vasilenko, F. G. Vashuk, I. F. Povkhan // *XX International Conference Promising ways and directions of improving the educational system, Uzhgorod, November 16–19, 2010. – Uzhgorod, 2010. – Vol. 21, № 1. – P. 217–222.*
22. Deng H. Bias of importance measures for multi-valued attributes and solutions / H. Deng, G. Runger, E. Tuv // *Proceedings of the 21st International Conference on Artificial Neural Networks (ICANN), Espoo, Finland, Jun 14–Jun 17, 2011. – Espoo, 2011. – P. 293–300.*
23. Kamiński B. A framework for sensitivity analysis of decision trees / B. Kamiński, M. Jakubczyk, P. Szufel // *Central European Journal of Operations Research. – 2017. – Vol. 26 (1) – P. 135–159.*
24. Dietterich T. G. An experimental comparison of three methods for constructing ensembles of decision trees: bagging, boosting, and randomization / T. G. Dietterich // *Machine learning. – 2000. – Vol. 40, № 2. – P. 139–157.*
25. Povhan I. Generation of elementary signs in the general scheme of the recognition system based on the logical tree / I. Povhan // *Collection of proceedings «Electronics and information technology». – 2019. – Vol. 12. – P. 20–29.*
26. Subbotin S. The dimensionality reduction methods based on computational intelligence in problems of object classification and diagnosis / S. Subbotin, A. Oliinyk // *Recent Advances in Systems, Control and Information Technology / eds. : R. Szewczyk, M. Kaliczyńska. – Cham : Springer, 2017. – P. 11–19. – (Advances in Intelligent Systems and Computing, vol. 543).*
27. Subbotin S. A. Methods and characteristics of localitypreserving transformations in the problems of computational intelligence / S. A. Subbotin // *Radio Electronics, Computer Science, Control. – 2014. – No. 1. – P. 120–128.*
28. Two-level clustering approach to training data instance selection: a case study for the steel industry / [H. Koskimaki, I. Juutilainen, P. Laurinen, J. Roning] // *Neural Networks : International Joint Conference (IJCNN-2008), Hong Kong, 1–8 June 2008 : proceedings. – Los Alamitos : IEEE, 2008. – P. 3044–3049. DOI: 10.1109/ijcnn.2008.4634228*
29. Subbotin S. The neuro-fuzzy network synthesis and simplification on precedents in problems of diagnosis and pattern recognition / S. Subbotin // *Optical Memory and Neural Networks (Information Optics). – 2013. – Vol. 22, № 2. – P. 97–103. DOI: 10.3103/s1060992x13020082*
30. Subbotin S. A. Methods of sampling based on exhaustive and evolutionary search / S. A. Subbotin // *Automatic Control and Computer Sciences. – 2013. – Vol. 47, № 3. – P. 113–121. DOI: 10.3103/s0146411613030073*
31. De Mántaras R. L. A distance-based attribute selection measure for decision tree induction / De Mántaras R. L. // *Machine learning. – 1991. – Vol. 6, № 1. – P. 81–92.*
32. Alpaydin E. *Introduction to Machine Learning* / E. Alpaydin. – London : The MIT Press. 2010. – 400 p.
33. Painsky A. Cross-validated variable selection in tree-based methods improves predictive performance / A. Painsky, S. Rosset // *IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2017. – Vol. 39, No. 11. – P. 2142–2153. DOI:10.1109/tpami.2016.2636831.*
34. Miyakawa M. Criteria for selecting a variable in the construction of efficient decision trees / M. Miyakawa // *IEEE Transactions on Computers. – 1989. – Vol. 38, No. 1. – P. 130–141.*