

Розглянуто підходи до класифікації факторів пожежної небезпеки з використанням методу дерева рішень. Проаналізовано найбільш поширенні алгоритми класифікації. За допомогою алгоритму C4.5 класифіковано основні та змінні фактори, що впливають на виникнення природних пожеж, та присвоєно їм вагові коефіцієнти. Запропоновано напрями подальшого використання отриманих результатів

Ключові слова: фактори пожежної небезпеки, інтелектуальний аналіз даних, алгоритм класифікації C4.5, дерево рішень

Рассмотрены подходы к классификации факторов пожарной опасности с использованием метода дерева решений. Проанализированы наиболее распространённые алгоритмы классификации. С помощью алгоритма C4.5 классифицированы основные и переменные факторы, влияющие на возникновение природных пожаров и им присвоены весовые коэффициенты. Предложены направления дальнейшего использования полученных результатов

Ключевые слова: факторы пожарной опасности, интеллектуальный анализ данных, алгоритм классификации C4.5, дерево решений

UDC 004.62:614.841.2.001.2

DOI: 10.15587/1729-4061.2016.79868

A DECISION TREE IN A CLASSIFICATION OF FIRE HAZARD FACTORS

N. Pashynska

PhD, Senior Researcher*

E-mail: n_pashynska@ukr.net

V. Snytyuk

Doctor of Engineering Science,
Professor, Head of Department*

E-mail: snytyuk@gmail.com

V. Putrenko

PhD, Senior Researcher
Education and scientific complex

"Institute for Applied System Analysis"

National Technical University of Ukraine

"Igor Sikorsky Kyiv Polytechnic Institute"

Peremohy ave., 37, Kyiv, Ukraine, 03056

E-mail: putrenko@wdc.org.ua

A. Musienko

PhD, Researcher*

*E-mail: mysienkoandrey@gmail.com

Department of intellectual and information systems

Taras Shevchenko National University of Kyiv

Volodymyrska str., 60, Kyiv, Ukraine, 01033

1. Introduction

At present, Ukraine experiences growing levels of natural and technogenic threats and fire hazards. In particular, every year there is a significant number of natural and technogenic disasters; among rather frequent risks are fires, including those of natural origin. In recent years, the problem of protecting forests from fire has become particularly acute because of higher air temperatures, rainfall shortage, and strong winds. Thus, in 2015 compared to 2014, the number of fires in natural ecosystems increased 2.2 times [1].

In general, Ukraine annually records on average about 3.5 thousand wildfires that destroy more than 5 thousand hectares of forest. In addition to forest fires, natural fires include those of peat, steppe, and farmland. They are of different nature, depending on the conditions of fire origin, vegetation, and soil. The main groups of natural factors include topography, vegetation, and climatic resources, which together determine the characteristics of a landscape structure and, to a certain extent and at a certain probability, contribute to fires. Therefore, the necessary steps in risk assessment are to determine the fixed and variable factors that affect the potential for fires and to classify them by the available features.

Such research is increasingly important because it is essential to classify the factors of fire danger and to define them with appropriate weight factors to improve the monitoring and prediction of natural fires in Ukraine.

2. Literature review and problem statement

In recent years, there has been an increase in the use of predictive analysis for studying and predicting natural fires. Active development of remote sensing (RS) and accumulation of obtained geospatial data facilitate an increase in the number of studies on the problems of monitoring natural fires by means of remote control [2]. However, the authors of this study focus their attention only on the technical features of the instrument MODIS. Common approaches to using GIS technology to assess fire threats and their mapping are considered in [3]. On the other hand, this study does not sufficiently consider analytical capabilities of geographic information tools for monitoring and predicting fires. The authors of [4] suggest using neural network algorithms to predict forest fires. The predictive information of this study is mainly based on meteorological parameters, without regard to other factors. The analysis of natural fires' dynamics

for several centuries through using six variables is described in [5], but the emphasis is placed on the characteristics of the terrain. The methods of spatial research analysis and mapping of fire threat areas are used by the authors of [6]. The focus is zonal factors (vegetation, soil cover, and underlying surface) beyond climatic factors. Statistical evaluation of risks of natural disasters, including fires, is explored in [7]. The study mainly examines the patterns of the distribution, occurrence and effects of natural fires. Spatiotemporal features of fires are investigated by the authors of [8], who have developed an algorithm of space-time prediction, using different methods of data mining with the main focus on the methodology of the study. Using the methods of multivariate regression data and artificial neural networks for predicting natural fires is analyzed in [9], which is based only on meteorological data. The problem of utilizing geospatial data to classify fire factors is considered in [10]. However, the classification contains only constant factors, without including variable (dynamic) factors. Methods of artificial intelligence are involved in the analysis of fires in [11]. The authors have developed a technique to study areas that endured natural fires, but they do not consider the factors of their occurrence. An evolutionary technology to determine the time and the spread of fire on the basis of production rules determined by experts is suggested by the authors of [12]. The modeling of the time and the spread of fire does not include the parameters of fire sources and their consequences.

The analysis of the above-mentioned approaches in the studies shows that they need considerable improvement. Thus, despite the large amount of studies on the use of predictive analysis for researching and predicting natural fires, it is necessary to specify the common risk assessment algorithms, depending on the environmental conditions of a territory. In particular, for estimating the models of the risk of forest fires, it is important to take into account classifications of fire factors.

3. The purpose and objectives of the study

The purpose of this study is to develop a classification of factors of fire danger with the use of a decision tree. According to the set goal, it is necessary to solve the following main tasks:

- to select and prepare geospatial data for the study;
- to choose an algorithm for the classification with the use of a decision tree;
- to classify fixed and variable factors affecting fire safety of an area;
- to assign weight coefficients to the main factors.

4. Data to study fire hazard factors

The study involved the use of geospatial data of the European Space Agency, which in recent years has expanded the group of space satellites and remote sensing and launched a program Copernicus. This program provides with free satellite imagery data and derived thematic products that can be used in the analysis of vegetation, climate change and safety of water resources, which are traditionally considered in determining the level of fire danger in an area.

Copernicus is a European system for monitoring the Earth, for which data are collected from various sources,

including Earth observation satellites and on-site sensors. They are further processed to provide accurate and relevant information in six thematic areas: land cover, sea waters, the atmosphere, climate change, management of emergency, and security [13].

Global terrestrial service coverage (Global Land Service) is part of the service of the Copernicus land cover, which provides a range of biogeophysical data on the status and evolution of the Earth's surface on a global scale at middle and low spatial resolutions. These data are used to monitor vegetation, water cycle, and energy balance. The thematic products of Copernicus can be used to study the dynamics of regional climate changes and natural threats by researching the dynamic and structural component of assessing the vegetation cover of Ukraine.

One of the areas that is also supported by the system is the accumulation of data on territories that have endured fires. The Burnt Area product contains data on land areas that were burned during a season. Its use helps both obtain integrated data for a certain period of time on the scope of the fires and determine the outbreak of new fires during the time interval between recordings.

These samples can be used as training data to create models of fire risk using other Copernicus products, which are collected by the same methodology.

Among the Copernicus products that can characterize a fire hazard, it is necessary to single out an assessment index NDVI (the Normalized Difference Vegetation Index), which is normalized relative to the vegetation index and represented as a simple quantitative indicator of the amount of photosynthetically active biomass (usually called the vegetation index). On the one hand, the NDVI can serve as an indicator of vegetation while, on the other hand, it reveals the presence of dry organic matter that is suitable for burning. This is one of the most common and used codes to solve problems on the basis of a quantitative assessment of vegetation. It is calculated by the following formula:

$$NDVI = \frac{NIR - RED}{NIR + RED},$$

where NIR is reflection in the near infrared spectrum, and RED is reflection in the red spectrum.

Accordingly, the density of vegetation (NDVI) at some point in the image equals the difference of the intensities of the reflected light in the red and infrared ranges, divided by the sum of their intensities.

Low NDVI values indicate a poor condition of vegetation, which can be caused by a drought and can consequently lead to an increased fire danger. In this case, it is important to maintain appropriate scale and resolution of data, which should be more generalized than for conventional agricultural monitoring.

In addition to the NDVI, an important index characterizing the fire condition of vegetation is the index of DMP, or performance of dry matter, showing an overall growth rate or dried increase in the vegetation biomass, which is expressed in kilograms of dry matter per hectare per day (kgDM/ha/day). The DMP is directly related to the NPP (net primary productivity), but its units are configured for agro-statistical purposes. In this case, high values of the DMP may contribute to fire risks within an area.

Another important measure is the groundwater index (the Standardized Water Level Index, SWI), which assesses the state of water at different depths in the soil. This index

indicates the occurrence of a drought and an increase in the fire hazard. It is mainly caused by deposition through a process of infiltration. The moisture content of the soil is a very heterogeneous variable, subject to small scale changes in the properties of soil samples and drainage. Satellite measurements are integrated on relatively large scale areas with presence of vegetation, which complicates interpretation. High values of the index reduce fire risk, but low values indicate a risk of increased danger of natural fires.

All parameters of natural systems are recorded with the time distinction of once every 10 days, so a prediction of a fire danger should be rendered once every 10 days, too.

A combined analysis of these data is used for determining integrated indicators that reflect a potential fire hazard area. For this purpose, a number of procedures of GIS analysis are used within the module Spatial Analyst ArcGIS, including reclassification functions and raster algebra.

5. A rationale for the chosen method of the study

To solve the problem of classifying numerous factors of fire, we suggest using the method of building a decision tree – a method of representing the rules hierarchically, by a consistent structure, where each object corresponds to a single node that produces the decision.

At first, decision trees were suggested by Hoveland and Hunt in the late 1950s.

The earliest known work to describe the nature of decision trees – “Experiments in Induction” – was published in 1966 [14]. The decision tree is a way of presenting rules in a hierarchical consistent structure. The basic component of this structure is the answer “Yes” or “No” if the tree is binary (the CART algorithm), or a series of questions identified dynamically when trees have an arbitrary number of branches (the C4.5 algorithm).

The model, which is presented in the form of a decision tree, is intuitive and conducive to understanding the problem being solved. This feature of decision trees is important not only in referring a new object to a class, but it is useful in interpreting the classification model as a whole. The decision tree makes it possible to understand and explain why a particular object belongs to a particular class.

Decision trees can create classification models in areas where it is difficult for analytics to formalize knowledge. The algorithm for constructing a decision tree does not require the user to select input attributes (independent variables). The input of the algorithm can admit all existing attributes; the algorithm itself will choose the most important among them, and only they will be used to build the tree.

The accuracy of the models created with the help of decision trees is comparable to other methods of building classification models (statistical methods and neural networks). By now, a number of scalable algorithms have been developed and used to build decision trees of very large databases. Most algorithms for constructing decision trees are capable of some special handling of missing values. Many classical statistical methods to solve classification problems can only work with numeric data, whereas decision trees can process both numerical and categorical types of data [14].

For automatic building of decision trees through training with examples, there have been developed a number of algorithms. They include, among others, CART, C4.5, NewId, ITrule, CHAID, and CN2.

The algorithm CART (Classification And Regression Trees) is designed to build a binary decision tree. Binary trees are also called binaries because each node in the tree partition has only two off-springs. This algorithm solves the problem of classification and regression.

The ID3 (Iterative Dichotomizer) algorithm is used to process natural language domains. ID3 is difficult to use for processing continuous data. If the value of any attribute is continuous, there are many opportunities of dividing information by this attribute. The ID3 algorithm works recursively, with the chosen feature being used in each node to break a set of data into subsets, based on the tree root that contains all the data.

The ID3 algorithm starts with the original data set as the root node. Each iteration of the algorithm determines an unused attribute set, and the entropy of this attribute is calculated, followed by the selection of an attribute that has the lowest entropy (or the greatest information gain) value. Next, the set is split by the selected attribute to obtain subsets of data. If all elements in the subset belong to the same class, this subset is not processed further. This node in the decision tree becomes terminal. The work of the ID3 algorithm ends if each subset has been classified. This algorithm generally produces small trees, but it does not always give the lowest possible tree [15].

A drawback of the ID3 algorithm is its incorrect processing of attributes with unique values for all objects of a training set. For such objects, information entropy is zero, and no new data can be obtained from the built tree by this dependent variable. The subsets that have been obtained after splitting will contain one object each. The C4.5 algorithm solves this problem by introducing normalization.

The algorithm C4.5, based on the algorithm ID3, adds a function of transferring the decision tree into equivalent rules and decisions for important research tasks. C4.5 adapts the information entropy method and selects an attribute for a maximum speed of information retrieval as well as a threshold segment as the best attribute of testing and a threshold segment.

The algorithm C4.5 simulates a decision tree with an unlimited number of branches in the node. This algorithm can only work with a discrete dependent attribute and, therefore, can only solve classification tasks. C4.5 is considered to be one of the best known and most used algorithms for building classification trees.

Working with the C4.5 algorithm, it is necessary to comply with the following requirements:

- each entry of a data set must be associated with a defined class, which means that one of the attributes of the data set should be the label of a class;
- the classes are to be discrete;
- each sample must clearly relate to one of the classes;
- the number of classes should be much less than the number of records in the test data set;
- the algorithm C4.5 works slowly with extra large and noised data sets.

A pseudocode of the C4.5 algorithm:

1. Input: database D
2. Tree = {}
3. If D is pure OR other stopping criteria then
4. Terminate
5. End If
6. For all attribute ϵ D do
7. Entropy for root

8. End For
9. best = Best attribute entropy
10. Tree = Create a node then test a best in the root
11. Dv = sud_datasets from D
12. For all Dv do
13. Tree v= C4.5(Dv)
14. End for
15. Return Tree [15]

6. The method of classifying fire hazard factors

The production of a decision tree:

1. The calculation of information entropy in the classification.

Suppose that S is the number of examples in the training sets where there are m classifications of samples C_i ($i=1, 2, \dots, m$). S_i is the number of samples in the classification C_i . The formula is as follows:

$$I(S_1, S_2, \dots, S_m) = \sum_{i=1}^m p_i \log_2(p_i),$$

where $p_i = S_i/S$ is the probability of a random examples belonging to C_i .

2. The calculation of information entropy of each attribute.

Let us assume that an attribute X has v values $\{x_1, x_2, \dots, x_v\}$ and divides S into v subsets $\{s_1, s_2, \dots, s_v\}$; S_j includes those examples with S that assume the value of x_j for the attribute X ($J=1, 2, \dots, v$). The expected entropy (entropy condition) of using the attribute X as a classification attribute is as follows:

$$E(X) = \sum_{j=1}^v \frac{S_{1j} + \dots + S_{mj}}{s} I(s_{1j}, \dots, s_{mj}),$$

where s_{ij} is the number of examples that relate to the classification C_i in the subset s_j , and

$$I(s_{1j}, s_{2j}, \dots, s_{mj}) = \sum_{i=1}^m p_{ij} \log_2(p_{ij}),$$

in which $p_{ij} = \frac{S_{ij}}{S_j}$ is the probability of each example in s_j belonging to C_i .

3. The calculation of the information gain, particularly regarding the attribute.

The information function of enhancing the X attributes:

$$\text{Gain}(X) = I(S_1, S_2, \dots, S_m) - E(X).$$

The function of the information gain tends to acquire great importance for testing the value that is likely to produce multiple branches. However, a test to produce multiple branches does not mean that the obtained predictive result will be the best for these unknown objects. The function of the information enhancement speed can compensate for the lack of information gain. The data gain transfer speed is an improvement in obtaining information, which can eliminate the influence of the attribute that produces multiple branches. The function of the information gain considers not only the number of units but also the size of each node (the number of examples) for each segment, and not only the amount of information included in the classification but also

each segment. The information gain of the X attributes is as follows:

$$A(X) = \frac{\text{Gain}(X)}{I(S_1, S_2, \dots, S_v)},$$

where v is the number of node branches, and S_i is the number of entries for the i-th branch.

4. The creation of a decision tree.

The information gain – Gain (X) – is calculated, followed by estimating the information gain speed A (X) for each attribute; for the test, the selected attribute is the one that has the greatest extent of information gain and the factor value of information gain, which is not lower than the average value for all attributes. In the test, the attribute is assumed to be a node, and each attribute distribution is perceived as a segment branch of examples. If all examples of a node belong to the same class, the node is a leaf, which is distinguished by its classification. It would be useful to form the initial decision about the tree by the method of recursion where all examples of each subset would receive the same value of the main attribute or an attribute for use when it is not in the sample [10].

7. The results of classifying fire hazard factors

7.1. Implementation of the decision tree for constant factors of fire danger within an area

Land cover	
Forest cover	Evergreen coniferous forest – 0.688
	Evergreen deciduous forest – 0.625
	Seasonal coniferous forest – 0.938
	Seasonal deciduous forest – 0.750
	Mixed forest – 0.625
Bushes	Thick bushes – 0.938
	Open shrubs – 0.875
Meadows	Forest steppe – 0.563
	Steppe – 0.500
	Grasses – 1.000
	Meadow wetlands – 0.063
Anthropogenically altered areas	Agricultural fields – 0.250
	Urban areas – 0.188
	Hayfields and perennial plants – 0.313
Relief	The angle of the surface slope, degree
	0–5 – 0.022
	6–10 – 0.067
	11–15 – 0.133
	16–20 – 0.222
	21–25 – 0.333
	26–30 – 0.467
	31–35 – 0.622
	36–40 – 0.800
	>40 – 1.000
	Exposure of the slope
	Northern (316–45°) – 0.100
	Western (226–315°) – 0.300
	Eastern (46–135°) – 0.600
	Southern (136–225°) – 1.000

Climatic factors	Height, m	0-483 – 1.000
		484-1,074 – 0.667
		1,075-1,709 – 0.200
		1,710-2,435 – 0.400
	Temperature, °C	0-3.2 – 0.067
		3.21-6.4 – 0.200
		6.41-9.59 – 0.400
		9.6-12.97 – 0.667
		>12.97 – 1.000
	Precipitation, mm	429-685 – 1.000
		686-861 – 0.667
		868-992 – 0.4
		993-1,180 – 0.2
	>1,181 – 0.067	

DPM	0-12 – 0.102	
	12-24 – 0.333	
	24-48 – 0.600	
	48-60 – 0.303	
	60-72 – 0.200	
	72-84 – 0.400	
	84-96 – 0.550	
	96-108 – 0.767	
	108-120 – 0.800	
	SWI	0-20 – 0.900
		20-40 – 0.850
		40-60 – 0.800
		60-80 – 0.700
		80-100 – 0.677
		100-120 – 0.500
		120-140 – 0.400
		140-160 – 0.333
	160-180 – 0.202	
	180-200 – 0.100	

Thus, the main groups of natural factors include land cover, topography, and climatic resources.

The determining factors of natural fires are the vegetation cover and the soil, the nature of which affects the ranking of potential danger of fires. Depending on the ratio between the extent of the soil cover and the vegetation types, it is possible to distinguish between various spatial groupings in terms of the risk of fires. The main features of vegetation according to the types are the following:

- forest cover, shrubs, or grass cover;
- seasonal or evergreen cover of forests, coniferous or deciduous forests;
- the ratio of vegetation types and the extent of the project soil cover.

For the soil cover, the important factors are open areas, soils' moisture and forming conditions – wetlands and peat deposits. There are important forms of land use, including agricultural territories of various types and urban areas.

The relief factor contains several important components, including the physical surface altitude, which determines insolation and dissected topography, surface slope, and slope exposure. The altitude affects the level of surface insolation and the degree of rainfall, humidity, and temperature.

Climatic factors also play a key role in assessing danger of natural fires. Traditionally, fire risk modeling takes into account a set of factors such as temperature, precipitation, moisture balance, solar radiation, as well as wind speed and direction. However, the sample was selected by only the basic factors: the average annual temperature in Celsius and the average annual rainfall in millimeters.

7. 2. Implementation of the decision tree for variables of a fire hazard within an area

NDI	0-25 – 0.050
	25-50 – 0.112
	50-75 – 0.200
	75-100 – 0.405
	100-125 – 0.600
	125-150 – 0.850
	150-175 – 0.600
	175-200 – 0.405
	200-225 – 0.333
	225-254 – 0.200

Thus, using a decision tree, we have classified the territory in terms of fire danger and obtained quantitative measurements that can act as weights for further modeling of fire hazards. The selected variable factors can have a nonlinear correlation with the fire hazard. For example, high and low values of NDI correspond to a low probability of fire, due to the low volume of biomass, and territories with index values that are average and below the average reflect a high probability of fire.

At the same time, low and high index values of DPM are more dangerous, whereas average values are of less impact. A high rate of groundwater index SWI corresponds to a low likelihood of fires.

8. Conclusions

We have determined a system of indicators of the ongoing fire risk factors (land cover, topography, and climatic resources) as well as variable indicators, which are taken into account while determining the level of a fire hazard in an area and are used to describe the state of vegetation, climate change, and supply of water resources. The study involves the use of geospatial data of Copernicus as a European system for monitoring the Earth, including the NDVI (normalized difference vegetation index), the DMP (dry matter productivity), and the SWI (standardized water level index).

1. To solve the problem of classifying fire factors, we suggest using the method of building decision trees, which is a method of arranging rules in a hierarchical consistent structure where each object corresponds to a single node through which the decision is made. The C4.5 algorithm is used to simulate a decision tree with an unlimited number of branches in the node and to represent causal relationships in the process of referring an object of classification to a particular class.

2. The use of the C4.5 algorithm makes it possible to build a decision tree and classify fixed and variable factors of fire danger. We distinguish between 3 main classes of permanent environmental factors, which include land cover, topography, and climatic resources. They, in turn, are divided into subclasses. In particular, in the forest cover subclass,

the greatest fire danger can arise in pine forests and thick bushes; in the relief subclass, it is important to take into account a surface slope exposure and the slope angle; in the subclass of climatic factors, the important indicators are high temperature and low rainfall. The construction of the decision tree for variable factors (the indices NDVI, DMP, and SWI) allow ranking these indicators according to their impact on the level of fire danger in an area.

3. The use of the C4.5 algorithm has helped obtain quantitative measurements that can act as weights for further

modeling of fire hazards. The resulting values range from 0 to 1, where a value of 0 prevents natural fires (e. g., water surfaces), but values close to 1 indicate a high hazard potential of natural fires.

4. The decision trees, obtained in the process of classifying, are important for planning measures to prevent natural fires. They can also be used for zoning in terms of fire hazards in spatial modeling of fires, mathematical modeling of their effects, as well as in further monitoring and prediction of natural fires.

References

1. Analitichnyi ohliad stanu tekhnogennoi ta pryrodnoi bezpeky v Ukraini za 2015 rik [Electronic resource]. – UkrNDITsZ, 2015. – Available at: <http://undicz.dsns.gov.ua/ua/Analitichniy-oglyad-stanu-tehnogennoi-ta-prirodnoi-bezpeki-v-Ukrayini.html>
2. Hawbaker, T. Detection rates of the MODIS active fire product in the United States [Text] / T. Hawbaker, V. Radeloff, A. Syphard, Z. Zhu, S. Stewart // *Remote Sensing of Environment*. – 2008. – Vol. 112, Issue 5. – P. 2656–2664. doi: 10.1016/j.rse.2007.12.008
3. Atlas of natural hazards & risks of Georgia [Electronic resource]. – Caucasus Environmental NGO Network. – Available at: <http://drm.cenn.org/index.php/en/> - 24.09.2016.
4. Yasynskyy, F. N. Prohnozovanye veroyatnosti vozniknovenyia lesnykh pozharov s pomoshch'yu neyrosetevoho alhorytma na mnohoprotsessornoy vychyslytel'noy tekhnike [Text] / F. N. Yasynskyy, O. V. Potëmkyina, S. H. Sydorov, A. V. Evseeva // *Vestnyk YHEU*. – 2011. – Vol. 2. – P. 1–4.
5. Oneal, C. Geographic analysis of natural fire rotation in the California redwood forest during the suppression era [Text] / C. Oneal, J. Stuart, S. Steinberg // *Fire Ecology*. – 2006. – Vol. 2, Issue 1. – P. 73–99. doi: 10.4996/fireecology.0201073
6. Jovanović, R. Spatial analysis and mapping of fire risk zones and vulnerability assessment – case study mt. Stara Planina [Text] / R. Jovanović, Ž. Bjeljac, O. Miljković, A. Terzić // *Journal of the Geographical Institute Jovan Cvijic, SASA*. – 2013. – Vol. 63, Issue 3. – P. 213–226. – Available at: <http://www.doiserbia.nb.rs/img/doi/0350-7599/2013/0350-75991303213J.pdf> doi: 10.2298/ijgi1303213j
7. Guo, H. Understanding global natural disasters and the role of earth observation [Text] / H. Guo // *International Journal of Digital earth*. – 2010. – Vol. 3, Issue 3. – P. 221–230. doi: 10.1080/17538947.2010.499662
8. Cheng, T. Applications of spatio-temporal data mining and knowledge for forest fire [Text] / T. Cheng, J. Wang // In: *Proceedings of the ISPRS Technical Commission VII Mid Term Symposium, 2006*. – P. 148–153.
9. Cortez, P. A Data Mining Approach to Predict Forest Fires using Meteorological Data [Text] / P. Cortez, A. Morais // *New trends in artificial intelligence: proceedings of the 13th Portuguese Conference on Artificial Intelligence (EPIA 2007), 2007*. – P. 512–523.
10. Putrenko, V. V. Data mining of the risk of natural fires based on geoinformation technologies [Text] / V. V. Putrenko // *Technology Audit and Production Reserves*. – 2016. – Vol. 4, Issue 3(30). – P. 67–72. doi: 10.15587/2312-8372.2016.76154
11. Özbayoğlu, A. Estimation of the Burned Area in Forest Fires Using Computational Intelligence Techniques [Text] / A. Özbayoğlu, R. Bozer // *Procedia Computer Science*. – 2012. – Vol. 12. – P. 282–287. doi: 10.1016/j.procs.2012.09.070
12. Snitjuk, V. E. Jevoľucionnoe modelirovanie processa rasprostraneniya pozhara [Text] / V. E. Snitjuk, A. A. Bychenko // *Proc. XIII-th Int. Conf. Knowledge-dialogue-Solution*. – 2007. – Vol. 6. – P. 247–254
13. Copernicus Global Land Service [Electronic resource]. – Available at: <http://land.copernicus.eu/global/products/dmp> – 24.09.2016 p.
14. Hunt, E. Experiments in induction [Text] / E. Hunt, J. Marin, P. Stone. - New York; London: Academic P, 1966. – 247 p.
15. Lepshova, E. S. Realyzatsiya y rasparallelyvanye alhorytma yntellektualnoho analyza dannykh, osnovannoho na dereviakh resheniy [Text] / E. S. Lepshova, V. A. Billih // *Vysokoproizvodytelnye paralelne vychysleniya na klasternykh systemakh, 2012*. – P. 247–251.
16. Quinlan, J. R. C4.5 : Programs for Machine Learning [Text] / J. R. Quinlan. – San Mateo : Morgan Kaufmann Publishers Inc., 1993. – 302 p.