НЕЙРОІНФОРМАТИКА ТА ІНТЕЛЕКТУАЛЬНІ СИСТЕМИ

NEUROINFORMATICS AND INTELLIGENT SYSTEMS

UDC 519.876.5:681.518.2

COMPUTATIONAL INTELLIGENCE METHODS TO PATIENTS STRATIFICATION IN THE MEDICAL MONITORING SYSTEMS

Bakumenko N. S. – PhD, Associate Professor of Theoretical and Applied Systems Engineering Department, V. N. Karazin Kharkiv National University, Kharkiv, Ukraine.

Strilets V. Y. – PhD, Associate Professor of Theoretical and Applied Systems Engineering Department, V. N. Karazin Kharkiv National University, Kharkiv, Ukraine.

Ugryumov M. L. – Dr.Sc., Professor of Theoretical and Applied Systems Engineering Department, V. N. Karazin Kharkiv National University, Kharkiv, Ukraine.

Zelenskyi R. O. – Assistant of the Department of Urology and Pediatric Urology, Kharkiv Medical Academy of Postgraduate Education, Kharkiv, Ukraine.

Ugryumova K. M. – PhD, Associate Professor, Department of Mathematical Modeling and Artificial Intelligence, National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine.

Starenkiy V. P. – Dr.Sc., Professor, Head of the Radiation Oncology Department, State Organization "Grigoriev Institute for Medical Radiology and Oncology of the National Academy of Medical Sciences of Ukraine", Kharkiv, Ukraine.

Artiukh S. V. – PhD, Senior Researcher of the Radiotherapy Group of the Radiology Department State Organization "Grigoriev Institute for Medical Radiology and Oncology of the National Academy of Medical Sciences of Ukraine", Assistant Professor of Radiology and Radiation Medicine, Kharkiv National Medical University, Kharkiv, Ukraine.

Nasonova A. M. – PhD, Clinical Oncologist, Head of the Clinical Oncology and Hematology Department, State Organization "Grigoriev Institute for Medical Radiology and Oncology of the National Academy of Medical Sciences of Ukraine", Kharkiv, Ukraine.

ABSTRACT

Context. In modern medical practice the automation and information technologies are increasingly being implemented for diagnosing diseases, monitoring the condition of patients, determining the treatment program, etc. Therefore, the development of new and improvement of existing methods of the patient stratification in the medical monitoring systems is timely and necessary.

Objective. The goal of intelligent diagnostics of patient's state in the medical monitoring systems – reducing the likelihood of adverse states based on the choice of an individual treatment program:

− reducing the probability of incorrectly determining the state of the patients when monitoring patients;

− obtaining stable effective estimates of unknown values of treatment actions for patients (corresponding to the found state);

− the choice of a rational individual treatment program for the patients, identified on the basis of the forecasted state.

Method. Proposed methodology, which includes the following computational intelligence methods to patient's stratification in the medical monitoring systems:

1) method of cluster analysis based on the agent-based approach – the determination of the possible number of patient's states using controlled variables of state;

2) method of robust metamodels development by means artificial neuron networks under a priori data uncertainty (only accuracy of measurements is known) in the monitoring data: a) a multidimensional logistic regression model in the form of analytical dependences of the posterior probabilities of different states of the patients on the control and controlled variables of state; b) a multidimensional diagnostic model in the form of analytical dependences of the objective functions (quality criteria of the patient's state) on the control and controlled variables of state;

3) method of estimating informativeness controlled variables of state at a priori data uncertainty;

DOI 10.15588/1607-3274-2023-1-3

4) method of robust multidimensional models development for the patient's state control under a priori data uncertainty in the monitoring data in the form of analytical dependencies predicted from the measured values of the control and controlled variables of state in the monitoring process;

5) method of reducing the controlled state variables space dimension based on the analysis of the variables informativeness of the robust multidimensional models for the patient's state control;

6) method of patient's states determination based on the classification problem solution with the values of the control and forecasted controlled variables of state with using the probabilistic neural networks;

7) method of synthesis the rational individual patient's treatment program in the medical monitoring system, for the state identified on the basis of the forecast.

Proposed the structure of the model for choosing the rational individual patient's treatment program based on IT Data Stream Mining, which implements the «Big Data for Better Outcomes» concept.

Results. The developed advanced computational intelligence methods for forecast states were used in choosing the tactics of treating patients, to forecast treatment complications and assess the patient's curability before and during special treatment.

Conclusions. Experience in the implementation of "Big Data for Better Outcomes" concept for the solution of the problem of computational models for new patient stratification strategies is presented. Advanced methodology, computational methods for a patient stratification in the medical monitoring systems and applied information technology realizing them have been developed. The developed methods for forecast states can be used in choosing the tactics of treating patients, to forecast treatment complications and assess the patient's curability before and during special treatment.

KEYWORDS: Information Technology Data Stream Mining, Medical Monitoring Systems, Machine Learning Methods, Mathematical Models and Methods for Patient Stratification.

ABBREVIATIONS

ANN is an artificial neural network;
DM is a decision-makers;
CDSS is a computer decision support system;
IDS is an intellectual diagnostics of systems;
MFFN is a multilayer feedforward artificial neural network;
MMS is a medical monitoring system;
RBF is a radial-basis function neural network.

NOMENCLATURE

$\vec{X}^{(0)}$ is an input vector of state parameters values;
$\vec{F}$ is a vector of state's target functions;
P is a number of the elements in the dataset;
H_0 is a dimension of the input vector;
H_{K+1} is a dimension of the output vector;
$\vec{F}^{(K+1)}$ is an output vector of the ANN model;
K is a number of neuron in the hidden layer;
X_j is a set of the elements from j^{th} cluster;
$\vec{x}_{jp}$ is p^{th} vector of the element in the j^{th} cluster;
P_j is a number of elements in the j^{th} cluster;
k^* is a number of clusters;
$d(\vec{x}_{jp}, \vec{c}_j)$ is an intra-cluster distance;
$d_1(\vec{x}_{jp}, \vec{c}_j)$ is Manhattan distance;
MD^2 is Mahalanobis distance;
D_{KL} is Kullback-Leibler entropy;
w_{jp} is a membership matrix;
$\vec{c}_j$ is a vector of cluster's centers;
$M(\vec{c}_j)$ is an average measure of the intra-cluster distance;
$LF(x)$ is a loss function;
Z is a target set of pairs of cluster's number and number of elements in them;
$\hat{Z}$ is a solution of an optimize problem;
k_j is the cluster's number;
D_X is correctness set;
$\rho(\vec{x}_{jp}, \vec{c}_j)$ is Cauchy distribution;
η is a parameter of Cauchy distribution;
n is an iteration number;
N is total number of iterations;
$k_t^{(n)}$ is estimated number of clusters on the n^{th} iteration;
f_i is selection function's values at the ANN output;
x_h is parameter's value at the input layer;
MV is selection function of the MV-problem;
ξ is significance level;
f_{fit} is a fitness function;
f_i^* is required value of the selection function;
$\sigma^0_{f_i,p}$ is standard deviation of the selection function;
β_f is regularization parameter;
n_α is number of measurements for fixed state;
$\chi^2_{f_i,p}$ is chi-squared test;
d_i is metric of the state's space;
L_{fit} is convergence acceleration factor;
$\mu_i(f_i^*)$ is membership functions;
$f_i(\vec{X}^{(0)})$ is fitness function value;
$\sigma^0_{f_i}$ is standard deviation of variables value f_i;
$f^*_{i,p}$ is required values of the function for p^{th} training set;
f^θ is normalized values of a variables or target functions;
f is current values of a variables or target functions;
$f_{\min}$ is minimum value of a variables or target functions;
$f_{\max}$ is maximum value of a variables or target functions;
l_f is slope factor of the activation function;

DOI 10.15588/1607-3274-2023-1-3

bt is slope factor of the function $th()$;

$\sigma^*_{X_h^{(0)}}$ is required standard deviation of the input data;

$X_{h.\max}^{(0)}$ is maximum value of the input data variables;

$X_{h.\min}^{(0)}$ is minimum value of the input data variables;

D_L is training dataset;

L is dimension of the training dataset (state's space);

$\hat{M}$ is search variables set;

M is variables subset from D_L set;

D_{F_i} is dispersion of the target function, in decibels;

$D_{F_i}^{(0)}$ is dispersion of the base (linear) model;

$D_{F_i}^{(model)}$ is dispersion of the other model;

$D_{M,\lambda}$ is subset of informativeness variables;

Σ_X is the covariance matrix of variables X_l and X_n;

σ_{X_l} is the standard deviation of X_l;

r_{ln} are the correlation coefficients between X_l and X_n (l=1...L, n=1...L);

E_i is a signal energy;

λ_{ih} is the informativeness coefficients;

D_Q is monitoring results;

Q is dimension of monitoring data;

Π^0 is design and regime parameters;

U^0 is control variables;

Φ^0 is phase variables;

t is forecast moment;

T_1 is lower forecast moment;

T_2 is upper forecast moment;

ε_t^0 is relative error;

Ψ is control state model;

Ω is diagnostic model;

W^0 is quality criteria of the patient's condition;

I is dimension of time series q^0;

$\vec{X}^*$ is vector of observed symptoms of unidentified precedent;

$X_1^*,...,X_A^*$ are principal components projections of $\vec{X}^*$;

A is number of the principal components;

c_{lk} are the classes centers (l, k=1…A);

$\rho(R_{k^*})$ is the probability distribution density agreeing to the Student's t-law;

$\left\langle \vec{X}_m \right\rangle$ is average projections value principal components vector of the observed symptoms of layer samples element;

$\Gamma(\ )$ is gamma function;

K_l is number of precedents in l^{th} class, l=1…A;

K_k is number of precedents in k^{th} class, k=1…A;

$\sum_{pooled}$ is the combined correlation matrix for the considered scenarios (for classes);

$P(R_k)$ is a prior probability of the classes realization;

$\hat{X}$ is quasisolution of MV-problem;

t_f is Student's coefficient for considered function f;

Ro_f is Romanovsky's coefficient for considered function f;

$M_\alpha[\ ]$ is mathematical expectation with significance level α;

x_p^* is required variables values x_p for prototype;

σ_p^* is required mean deviation values of variables x_p for prototype;

σ_{x_p} is mean square deviation of variables $x_p \in X^0$;

$\sigma_{f_i}^*$ is the mean deviation values of decision making criterion f_i for prototype;

σ_{f_i} is mean square deviation of decision making criterion $f_i \in F$;

γ is a regularization parameter;

time is the value of the survival time for the patient after undergoing treatment.

INTRODUCTION

The actual problems of modern medicine – the problems of assessing the states, forecasting the outcome of diseases, the effectiveness of treatment methods, assessing the likelihood of complications in patients – can be solved based on the use of advanced machine learning methods and information technologies that implement them. That's why developing an applied information technology for patient stratification in medical monitoring systems based on advanced machine learning methods is an actual scientific and practical task.

An MMS consists of monitoring hardware for the patient's condition, decision makers (physicians) and CDSS. The aim of such systems – to ensure continuous observation, information collecting, data processing and analyzing patient's condition, forming recommendations for treating [1, 2].

National Aerospace University named by N. Zhukovsky "Kharkiv Aviation Institute" for more than 15 years has been developing applied information technology for patient stratification in medical monitoring systems based on machine learning methods. The university has developed its own software to accomplish the assigned problems and has qualified personnel who participate in international projects in the field of applied mathematics, statistics, and machine learning. In cooperation with the Kharkov National Medical University, the developed software is being tested for stratification of patients with prostate cancer, squamous cell carcinoma of the head and neck [3–14]. The developed methods for forecasted states are also planned to be verified on the data obtained in various oncological pathologies and to use them in choosing the tactics of treating patients, to forecast treatment complications and assess the patient's curability before and during special treatment.

The material and technical support and personnel base of the Kharkov National Medical University allows col-

DOI 10.15588/1607-3274-2023-1-3

lecting data on the course of treatment of cancer patients receiving chemoradiation treatment, using the whole range of non-invasive diagnostic methods, a wide range of tumor markers. Working with patients is regulated by the Bioethics and Deontology Committee and complies with international GCP standards.

As part of the implementation of projects, an innovative strategy is proposed for choosing a rational individual tactics for treating patients based on forecasted states identified using robust stratification methods, which will improve clinical results and prevent complications.

It is also planned to collect data on patients with cancer who have undergone COVID-19 before starting treatment and to assess the effect of the latter on the timing of the start of radiation and chemotherapy, depending on the severity of post Covid's syndrome.

This paper shows computational models for processing and research about a patient's condition and his body functioning based on the "Big data for better outcomes" concept for MMS.

The object of study is the processes for predicting the states and choosing individual treatment programs for patients based on monitoring data.

The subject of study is the computational intelligence methods to patient's stratification in the medical monitoring systems.

The purpose of the work is to reduce the likelihood of adverse states for patients in medical monitoring systems based on the choice of an individual treatment program by means of computational intelligence methods.

1 PROBLEM STATEMENT

Reducing material costs and time of forecasting of the patients' states in the medical monitoring systems are possible through the automation forecasting process of the patient's states using information technology Big Data. We define Big Data as an information technology based on the use of approaches series, tools and processing methods of the structured and unstructured data of large volumes and considerable variety to obtain the results for DM that are effective in conditions of continuous increase and distribution by numerous nodes of the computer network data stream.

The goals of intelligent diagnostics of patient's state in the medical monitoring systems:

– reducing the likelihood of adverse states based on the choice of an individual treatment program;

– reducing the probability of incorrectly determining the state of the patients (errors of the third kind in classifying the state of the systems) when monitoring patients;

– obtaining stable effective estimates of unknown values of treatment actions for patients (corresponding to the found state);

– the choice of a rational individual treatment program for the patients, identified on the basis of the forecasted state.

The aim of the investigation is to develop advanced methodology for solving the synthesis problem of patient's individual treatment program in the medical monitoring systems by means of computational intelligence methods for the medical information analysis and applied information technology realizing them by means of new robust estimation methodology (M-estimation) based on the concept of invariance of the theory of optimal control and apply it to solving nonlinear multidisciplinary problem of under uncertainty.

Let us input data $\left(\vec{X}^{(0)}, \vec{F}\right)_p, p = 1...P$ are the results of the anamnesis (personal, visual examination, laboratory data), these are statistical data (dataset) from the observed patients, accumulated by medical institutions, as well as the individual results of patient monitoring, considered as training pairs. An input vector and output vector is of dimension H_0 and H_{K+1} respectively.

The data model, such as information from the ambulatory card, formed by experts (physicians).

Statistical data (dataset) from the observed patients had transferred to the database storage Big Data from medical organizations.

Each patient must be able to contribute their own monitoring results to the database.

Data processing and diagnostic results transmission carried out by using Internet resources, which is available to every user, including the use of CDSS on remote servers.

Structuring new mathematical statements and developing computational methods is necessary for solving the synthesis problem of a patient's individual treatment program in the medical monitoring systems by means of computational intelligence methods, in stochastic formulation (MV-problem).

On the result of processing the input data, you need to find the following:

– the required number of controlled state variables for the classification of states (single or multiple diseases, and their corresponding disease stages);

– results of the patient's condition classification (single or multiple cases and their corresponding disease stages);

– estimates of survival rates and clinical effectiveness of the treatment;

– quality of life assessments for patients during treatment and in the follow-up period;

– results of informative monitored condition variables synthesis for patients individually at the current point in the monitoring process;

– results of predicting the time series of monitored state variables (including medical influences);

– justification of the time intervals choice for measuring values of controlled variables during monitoring depending on the disease stage;

– the results of predicting patient conditions that cannot be assessed using current clinical, laboratory and in-

DOI 10.15588/1607-3274-2023-1-3

strumental methods (e.g. hormone resistant prostate cancer);

– selection of an individual treatment regimen (medical interventions) depending on the patient's condition in order to minimize the risk of serious adverse events.

Such developments compared to existing methods will ensure a reduction in the probability of incorrectly determining the state of the systems (errors of the third kind in classifying the state of the patients), as well as obtaining stable effective estimations of the unknown values variables (corresponding to the found state).

2 REVIEW OF THE LITERATURE

Over the past decade, as a result of cooperation, the authors have developed the methods for solving such problems:

– formation of a subset of controlled variables of state, the values of the quantities of which are check in by measuring instruments [3, 9];

– patient's classification of state results [4, 6, 7, 9];

– informativeness evaluations of the variables of for different stages of patients' diseases [4, 5, 9];

– robust metamodels: multidimensional diagnostic model, multidimensional logistic regression [6, 8, 9, 11];

– robust multidimensional models of control the states of the patients, evaluation of forecasted values of the controlled variables of state based on medical monitoring data for patients [11, 5];

– cluster analysis results – the number of recognizable states [5, 9, 12];

– patient's state classification results, using the values of the control and forecasted controlled variables of state [9, 11];

– the results of the synthesis of a rational individual treatment program for the patients for the state, determined on the basis of a forecast (determination of control variables (medical actions), that ensure the implementation of the treatment program) [9, 11, 14].

Numerical research was carried out with the help of the computer program "Non-linear evaluation methods in multicriteria problem of robust optimal designing and diagnostics of systems under parametric a priori uncertainty (methodology, methods, techniques and computer systems of support and decision-making implementing them)" (ROD&IDS) [13] developed by the authors.

The volumes of medical information about various diseases and their course are extremely large, and machine learning methods make it possible to process the accumulated information, take into account millions of different factors, social, territorial, demographic, genomic, etc., and make it possible to identify the unique features of each patient [15, 16]. It has been proven that the automation of the collection processes and further analysis of medical data allows to increase the accuracy of early diagnosis, prediction of the disease's development and the treatment effectiveness assessment [17]. For example, the Frost & Sullivan agency notes that artificial intelligence technologies increase the accuracy of diagnoses by 30–40%, and pathologist Andy Beck from Harvard Medical School believes that the further use of artificial intelligence technologies will reduce the errors rate in diagnosis by 85% [18].

Modern technologies are also used to choose the most effective treatment strategy. Recently, scales (systems) for objectifying the assessment of clinical-physiological and laboratory parameters have been used to choose treatment tactics, the scope of anesthetic support and surgical intervention, predict the frequency of probable postoperative complications, lethality, and the treatment effectiveness [19]. The priority of such a strategy is that after receiving the sum of objective indicators, the doctor converts them into a score, which is ranked into numerical and staging corridors. This makes it possible to comprehensively assess the patient's condition at the moment, monitor his condition and carry out appropriate treatment. The correctness and effectiveness of treatment with this method depends on the qualifications and experience of the doctor treating the patient [20].

The choice of mathematical methods for the description and research of biological and medical objects depends both on the specialist's individual knowledge and on the specifics of the tasks to be solved [21].

The following methods are most often used to solve the classification problem of the medical facilities state:

– binary classification (decision trees and random forests) [22];

– artificial neural networks (ANN) [23];

– multidimensional logistic regression [24];

– naive Bayesian classifier [25];

– support vector machines [23, 26].

Obviously, the use of each of these methods separately does not allow solving the general problem of synthesizing an individual patient treatment program in a medical monitoring system.

So, there is a need to structure the system model of medical decision-making as a sequence of interrelated tasks and the corresponding system model of decision synthesis. In other words, it is necessary to create a methodology for the synthesis of solution to the problem of making medical decisions in general.

The work proposes methodology for solving the synthesis problem of a patient's individual treatment program in the medical monitoring systems by means of computational intelligence methods for the medical information analysis.

3 MATERIALS AND METHODS

Medical Monitoring System is a set of monitoring states hardware, information technology Big Data tools (which contains CDSS), patients and decision-makers (doctors) who are in a communicate relation with each other and united with the purpose of managing and organizing the process of systematic or continuous observation, collection, processing and information investigation about the object state (of patient), its functioning (of various organs) and development for a certain period of time.

DOI 10.15588/1607-3274-2023-1-3

Its system is created and regulated by the monitoring entities (physicians) to ensure full, timely and accurate information and appropriate organization of effective functioning and control of the object of diagnosis (patient).

The Context diagram of "Big Data for Better Outcomes" concept implementation in the medical monitoring systems is shown in Fig. 1.

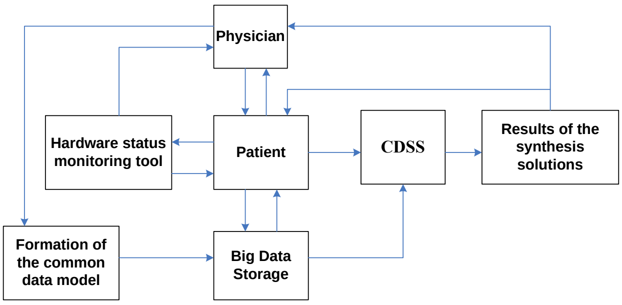

Figure 1 – Context diagram of "Big Data for Better Outcomes" concept implementation in the medical monitoring systems

Generalized methodology for solving the problem of synthesis problem of patient's individual treatment program in the medical monitoring systems by means of computational intelligence methods as a result of its decomposition may be given as a sequence of processing techniques of structured and unstructured data of large volumes and considerable diversity using developed software. The solving result is sequence of steps to obtain mathematical model $\vec{F}^{(K+1)}\left(\vec{X}^{(0)}\right)$ in order to solve diagnosing problems.

Patient stratification is carried out in two steps: Data preprocessing and Monitoring the state of the patients. Each stage is described below.

1. Data Preprocessing (Preliminary data preparation is performed by a doctor):

1.1 Formation of a subset of controlled variables of state $\left(\vec{X}^{(0)},\vec{F}\right)_p, p=1...P$, the values of the quantities of which are checked in by measuring instruments.

1.2 A total dataset of alternatives is generated. Each alternative includes subsets of control variables and controlled variables of state; objective functions. The total dataset includes subsets of alternatives corresponding to different states of the patients, including the healthy state.

1.3 Data cleaning from anomalous values of the quantities (outliers). Normalization of data.

1.4 Cluster analysis – the determination of the possible number of states (states of the patients) using control and controlled variables of state. We are looking for an answer to the question whether there is enough data from medical research to be able to recognize different states and the corresponding stages of diseases with the maximum a posteriori probability? If the number of identified clusters coincides with the specified one and the distances between the clusters are statistically significant, then the subset of monitored variables of state can be considered complete. Otherwise, the diagnostic system, which is used for monitoring of patient's states, should be equipped with new measuring instruments.

Based on the agent-based approach and in accordance with the chosen measures of intra-cluster distance, permissive elite selection rules are proposed for the formation of clusters, the selection of the best of them, and also for the selection of elements into clusters in the process of solution synthesis.

The result of solving such a problem is the number of clusters, as well as the number of elements in them.

The c-means clustering method was chosen as the basis.

Let the sample of data be considered as $X=\{X_j\}, X_j=\{\vec{x}_{jp}\}$ and $P=\sum_{j=1}^{k^*} P_j^*$ is the total number of elements.

It is necessary to find $\{k_j, P_j\}, j=1...k^*$.

Four measures of intra-cluster distance were used as a metric for the clustering data:

$$d\left(\vec{x}_{jp},\vec{c}_j\right)=\begin{cases} d_1\left(\vec{x}_{jp},\vec{c}_j\right),\\ MD^2\left(\vec{x}_{jp},\vec{c}_j\right),\\ w_{jp}^{-1}MD^2\left(\vec{x}_{jp},\vec{c}_j\right),\\ -D_{KL}\left(\vec{x}_{jp},\vec{c}_j\right).\end{cases}$$

Let us define the average measure of the intra-cluster distance:

$$M\left(\vec{c}_j\right)=\frac{1}{P_j}\sum_{p=1}^{P_j} d\left(\vec{x}_{jp},\vec{c}_j\right).$$

Also the loss function is defined as

$$LF(X)=\frac{1}{k^*}\sum_{j=1}^{k^*} M\left(\vec{c}_j\right).$$

Then the research problem statement will take in the form:

$$\begin{cases} Z=\left(k_j,\{P_j\}\right),\\ \hat{Z}=\arg\min_{X\in D_x} LF(X).\end{cases}$$

It is necessary to determine the number of clusters and distribute the data among clusters so that the value of the loss function is minimal.

To correct the centers of clusters, we use the expression

$$\vec{c}_j=\sum_{p=1}^{P_j} w_{jp}\vec{x}_{jp} \Big/ \sum_{p=1}^{P_j} w_{jp},$$

DOI 10.15588/1607-3274-2023-1-3

where $w_{jp} = \dfrac{\rho(\vec{x}_{jp}, \vec{c}_j)}{\sum_{p=1}^{P_j} \rho(\vec{x}_{jp}, \vec{c}_j)}$ is the membership matrix with the Cauchy distribution

$$\rho(\vec{x}_{jp}, \vec{c}_j) = \frac{1}{\pi\eta\left[1 + \dfrac{MD^2(\vec{x}_{jp}, \vec{c}_j)}{\eta^2}\right]}.$$

Data clustering algorithm based on agent-based approach:

1) set $k_t^{(n)} > k^*$, $P_j^{(n)} = \mathrm{int}\left(\dfrac{N}{k_t^{(n)}}\right)$ and randomly generate cluster centers $\{\vec{c}_j\}$;

2) using the selected measure with cluster distance $d_1(\vec{x}_{jp}, \vec{c}_j)$, choose $P_j^{(n)}$ the nearest neighbors for each j^{th} cluster;

3) for each j^{th} cluster, using the $\{w_{jp}\}$ and $\{\rho(\vec{x}_{jp} | P_j)\}$, the cluster centers $\{\vec{c}_j\}$ are corrected;

4) for each j^{th} cluster, using the selected measure within the cluster distance $d(\vec{x}_{jp}, \vec{c}_j)$ and $P_j^{(n)}$ the nearest neighbors are chosen. Delete the points that are duplicated: $P_j^{(n)} \to P_j$;

5) for each j^{th} cluster the average measure of the intra-cluster distance $M(\vec{c}_j)$ is calculated, and also $LF(X)$ is calculated;

6) elite selection. Find the cluster with the largest $M(\vec{c}_j)$ and delete it;

7) $k_t^{(n+1)} = k_t^{(n)} - 1$; $P_j^{(n+1)} = \mathrm{int}\left(\dfrac{N}{k_t^{(n+1)}}\right)$;

8) back to step 2, if $k_t^{(n+1)} > 1$.

1.5 Development of robust metamodels at a priori data uncertainty in the monitoring data (given that the results of measurements of variables of states are random variables – only accuracy of measurements is known):

a) a multidimensional logistic regression in the form of analytical dependences of the posterior probabilities of different states of the patients on the control and controlled variables of state;

b) a multidimensional diagnostic model in the form of analytical dependences of the objective functions (quality criteria of the patient's state) on the control and controlled variables of state.

To development of robust metamodels as initial information used the vector function is given by a training sample $(\vec{X}^{(0)}, \vec{F})_p, p = 1...P$. We must approximate the given set. The problem can be solved with a resultant mathematical mechanism, which may give any value of the vector function $\vec{F}_p^{(K+1)}(\vec{X}_p^{(0)})$, represented by this training set at a fixed input vector within the range, limited by the input data.

A multilayer feedforward artificial neural network (MFFN) and radial basis function network (RBF), used for data approximation, is a parallel distributed processor, which is capable of saving acquired knowledge and processing information between local processor elements (neuro-elements or neurons), bound by special links (synaptic links).

To provide parameter stability (robustness) and informative capability of statistical systems and processes models on the basis of learning ANN at the a priori input data uncertainty and also practically sufficient data approximation, it is reasonable to use advanced deep learning methods – stable (robust) statistical assessment of their parameters with adaptive learning rate as the ANN learning method.

Student and V. I. Romanovsky are used as a smoothing functional when choosing a rational solution, which provides a stable (robust) estimation of the searching values with parametric uncertainty of the input data, as well as sufficient, from a practical point of view, accuracy of data approximation in problems of improving systems.

The function (MV-problem) was used as a scalar convolution of selection functions, considering $f_i \equiv F_i^{(2)}, x_h = X_h^{(0)}$:

$$MV = \frac{1}{2PI} \sum_{p=1}^{P} \xi^{P-p} \sum_{i=1}^{I} \left(f_{fit}\left[\left(\frac{\mu_i(f_{i,p}^*)\Delta_{f_i,p}}{f_i^*(1+\sigma_{f_i,p}^0)}\right)^2\right] + \beta_f \cdot f_{fit}\left(\frac{\left|\chi_{f_i,p}^2 - n_\alpha\right|}{\sqrt{2n_\alpha}}\right) \right),$$

here $I=H_{k+1}$, $\xi = [0.95, 0.99]$,

$$f_{fit}(d_i) = 1 - \exp\left[-\frac{L_{fit}}{4} d_i\right], L_{fit} \ge 4, (d_i>0);$$

$$\Delta_{f_i} = F_i^{(K+1)}(\vec{X}^{(0)}) - f_i(\vec{X}^{(0)}), \sigma_{f_i}^0 = \frac{\sigma_{f_i}}{\sigma_{f_i}^*},$$

$$\frac{\left|\chi_{f_i,p}^2 - n_\alpha + 3\right|}{\sqrt{2(n_\alpha - 3)}} = \frac{n_\alpha}{\sqrt{2(n_\alpha - 3)}} \left| (\sigma_{f_i}^*)^2 - 1 + \frac{3}{n_\alpha} \right|, n_\alpha = 4.$$

For fitness functions $f_i(\vec{X}^{(0)})$ in the expression for scalar convolution of selection functions MV the meanings of relative values are calculated by formulas:

DOI 10.15588/1607-3274-2023-1-3

– direct conversation $f^0 = \dfrac{2l_f\left(f - \langle f \rangle\right)}{f_{\max} - f_{\min}}$, here $\langle f \rangle = \dfrac{f_{\max} + f_{\min}}{2}$, $f^0 \in [-1,1]$;

– inverse conversation $f = \left[\left(f_{\max} - f_{\min}\right) f^0 / l_f + \left(f_{\max} + f_{\min}\right)\right] / 2$, here $l_f = th(bt)$ for MFFN and $l_f = 1$ for RBF.

The relative mean square deviations of input data are calculated by formula

$$\left(\sigma^0_{f_i}\right)^2 = \left(\frac{2l_f}{f_{i,\max} - f_{i,\min}}\right)^2 \left(\frac{\Delta^0_{f_i}}{300} f_{i,\max}\right)^2 n_\alpha ,$$

$\Delta^0_{f_i} = \dfrac{\Delta f_i}{f_{i,\max}} 100\%$, here $l_f = th(\beta)$ for MFFN and $l_f = 1$ for RBF;

$$\left(\sigma^*_{X_h^{(0)}}\right)^2 = \left(\frac{2}{X^{(0)}_{h.\max} - X^{(0)}_{h,\min}}\right)^2 \left(\frac{\Delta^0_{X_h^{(0)}}}{300} X^{(0)}_{h.\max}\right)^2 n_\alpha .$$

By known training set D_L the ANN parameters vector $\hat{M}$ will choose according to the principle of maximum aposteriori probability distribution density:

$$\hat{M}_{n+1} = \underset{M \in D_L}{\arg\min}\, MV\left(M \middle| D_L\right).$$

An advanced computational method is proposed for estimating the parameters of structural-parametric models of systems and processes in the form of trained ANNs based on the method of stochastic approximation and the use of a regularizing sequential (adaptive) algorithm for the synthesis of solutions with delayed correction (based on the ravine method of conjugate gradient methods and the method of simulating the movement of bee colonies), which implements adaptive control of calculations in accordance with the principle of minimum disturbance.

Application of the proposed methods avoids the appearance of false ravines or valleys on response surfaces in case of gross errors in the input data.

In the paired comparison of metamodels models changing of the signal variance is evaluating, which defines the robustness of a particular model:

$$D_{F_i}, dB = 10\lg\left(\frac{D^{(model)}_{F_i}}{D^{(0)}_{F_i}}\right),$$

here *model* = 1 for MFFN and *model* =2 for RBF.

1. Monitoring of the state of the patients (data processing and analysis is performed by the doctor together with the patient):

2.1 The values of the control and controlled variables of state corresponding to the patient's state at a given time are measured.

2.2 Determination of the state for which the maximum a posteriori probability of its realization corresponds with the observed values of the quantities based on the solution of the classification problem, which allows to determine the disease stages that are not recognized by modern biomarkers (e.g., hormone-resistant stage in prostate cancer).

2.3 Estimation informativeness controlled variables of state at a priori data uncertainty, the synthesis of the set of informative controlled variables of state according to the patient's state (disease stage) for the reduction of the variables of state space dimension, using multidimensional diagnostic model, i.e. searching informative subset $D_{M,\lambda}$ of minimal dimension where $D_{M,\lambda} \subset D_L$.

The set of input data $F_i(X)$, where $X = \{x_l\}$, $l = 1 \ldots L$ is presented as a Taylor series, while retaining only the terms of the first infinitesimal order in the expansion. For the dispersion of an arbitrary gotten linear function of several random variables estimate holds:

$$D_{F_i} = \left(grad\, F_i\right)^T \Sigma_X\, grad\, F_i = \ldots = \sum_{l=1}^{L} \left(\frac{\partial F_i}{\partial x_l}\right)^2 \sigma^2_{x_l} + \sum_{l=1}^{L} \sum_{n=1, n \neq l}^{L} r_{ln} \frac{\partial F_i}{\partial x_l} \frac{\partial F_i}{\partial x_n} \sigma_{x_l} \sigma_{x_n}.$$

Let us define the signal energy by the expression

$$E_i = \sum_{h=1}^{H_0} \left| D_{F_i^{(2)} \left| X_h^{(0)} \right.} \right|,$$

here

$$D_{F_i^{(2)} \left| X_h^{(0)} \right.} = \left(\frac{\partial F_i^{(2)}}{\partial X_h^{(0)}}\right)^2 \sigma^2_{X_h^{(0)}} + \left(\sum_{n=1, n \neq h}^{H_0} r_{ln} \frac{\partial F_i^{(2)}}{\partial X_n^{(0)}} \sigma_{X_n^{(0)}}\right) \frac{\partial F_i^{(2)}}{\partial X_h^{(0)}} \sigma_{X_h^{(0)}}.$$

The informativeness coefficients (contribution weight $X_h^{(0)}$ into $F_i^{(2)}$) are defined by

$$\lambda_{ih} = \frac{\left| D_{F_i^{(2)} \left| X_h^{(0)} \right.} \right|}{E_i}, \quad \sum_{l=1}^{L} \lambda_{ih} = 1 .$$

2.4 Development of robust multidimensional models of control the state for the patients at a priori data uncertainty in the monitoring data in the form of analytical dependencies predicted from the measured values of the control and controlled variables of state in the monitoring process.

2.5 Forecasting multidimensional time series of controlled variables of state based on multidimensional models of control of the state for the patients:

DOI 10.15588/1607-3274-2023-1-3

$$D_Q = \{q^0(t+1)\},$$
$$q^0(t+1) = \left[\Pi^0(t+1), U^0(t+1), \Phi^0(t+1)\right],$$

where $t=T_1...T_2$ is a limited set of Π^0, U^0, Φ^0 and creating time series.

It is required to obtain a functional dependence, that will be reflect relationship between the next and previous values of the time series which satisfies the system preferences of DM, for a given forecast horizon:

$$q^0(t+1) = F\left(q^0(t+T_2-1),...,q^0(t-T_1)\right) + \varepsilon_t^0 .$$

Controlled process mathematical model:

$$\Delta\Phi^0 = \Psi\left(\Delta\Phi^0, \Delta U^0\right),$$
$$\Delta W^0 = \Omega\left(\Delta\Phi^0, \Delta U^0\right),$$

where the first expression is recurrent mathematical model for monitoring and the second expression is a diagnostic model:

$$X^0(t+1) = ln\left(\frac{q_i(t+1)}{q_i(t-T_1)}\right), i = 1...I,\ t = (-T_1+1)...T_2 .$$

2.6 Reducing the dimension of the space of controlled variables of state based on the analysis of the informativeness of the variables of the robust multidimensional models of control the state for the patients (Sensitivity Analysis). Estimating the rank of time series cointegration.

2.7 Determination of the states of the patients based on the solution of the classification problem and the values of the control and forecasted controlled variables of state is conducted using probabilistic neural networks. An input layer $X_1^*,..., X_A^*$ elements are the principal component vector projections of the monitored state variables $\vec{X}^*$. The layer of samples $c_{1k},…,c_{Ak}$ are the classes centers of the training samples. The number of exemplars equals the number of classes in the training set. The layer of input and the samples are entirely meshed structure. An element activity of the samples layer was determined by the dependence that related to the probability distribution density agreeing to the Student's t-law (that is proper for the limited samples):

$$c_{lk} = \rho(R_k) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi n}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{t_{lk}^2}{n}\right)^{-\frac{n+1}{2}},$$

here Γ is gamma function with $n=K_l+K_k-2$;

$$t_{lk} = \sqrt{\frac{MD_{lk}^2}{\frac{1}{K_l}+\frac{1}{K_k}}},$$

here MD_{lk}^2 is Mahalanobis distance from an unidentified precedent (it is supposed that it appropriate to l class) to the k sample:

$$MD_{lk}^2 = \tfrac{1}{A}\left(\vec{X}^* - \left\langle \vec{X}_k \right\rangle\right)^T \Sigma_{pooled}^{-1}\left(\vec{X}^* - \left\langle \vec{X}_k \right\rangle\right),$$

here $\sum_{pooled}$ is determined because of belonging of the precedent to one, or another class:

$$\Sigma_{pooled} = \frac{1}{K_l+K_k-2}\left((K_l-1)\Sigma_l + (K_k-1)\Sigma_k\right).$$

The output layer k^*, $\rho(R_{k^*})$ is a discriminator of the threshold value indicating the class to which belongs an unknown precedent.

The values of t-Student statistics depend on the selection of the degree of proximity of the precedent to the samples (when solving the classification problem), as well as between the samples (when analyzing the significance of the distance between classes). Therefore, it is necessary to further forming of supplementary statistical decision rule selecting a single guide basis. According to the maximum likelihood rule as a criterion for the transition from one state to the other state of patients, Bayes formula is used:

$$\forall k = 0...A-1: \frac{\rho\left(\vec{X}_{k+1}^*\right)}{\rho\left(\vec{X}_k^*\right)} \ge 1 .$$

Which is true provided that

$$\frac{P(R_k)}{P(R_{k+1})}\cdot\frac{\rho\left(\vec{X}_{k+1}^*\right)}{\rho\left(\vec{X}_k^*\right)} \approx 1 .$$

2.8 Synthesis of a rational individual treatment program for the patients in the medical monitoring system, for the state identified on the basis of the forecast.

The control variables $\hat{X} = \left(M[X_0], \sigma_x^0\right)$ estimation problem can be presented like a MV-problem.

The MV-problem provides an efficient stable (robust) estimation of the required quantities under parametric input data uncertainty.

The quasisolution of this problem is

$$\hat{X} = \arg\inf_{\hat{X}\in D_X} MV\left(\hat{X}\middle| t_f, Ro_f\right),$$

where the scalar convolution of selection functions MV is

DOI 10.15588/1607-3274-2023-1-3

$$MV=\frac{1}{2I}\sum_{i=1}^{I}\left(\begin{array}{l} f_{fit}\left[\left(\dfrac{\mu_i\left(f_i^*\right)\Delta_{f_i}}{f_i^*\left(1+\sigma_{f_i}^0\right)}\right)^2\right]+\\ +\beta_f\cdot f_{fit}\left(\dfrac{\left|\chi_{f_i}^2-\left(n_\alpha-3\right)\right|}{\sqrt{2\left(n_\alpha-3\right)}}\right)\end{array}\right)+$$

$$+\gamma\frac{1}{2P}\sum_{p=1}^{P}\left(\begin{array}{l} f_{fit}\left[\left(\dfrac{\mu_p\left(x_p^*\right)\Delta_{x_p}}{x_p^*\left(1+\sigma_{x_p}^0\right)}\right)^2\right]+\\ +\beta_x\cdot f_{fit}\left(\dfrac{\left|\chi_{x_p}^2-\left(n_\alpha-3\right)\right|}{\sqrt{2\left(n_\alpha-3\right)}}\right)\end{array}\right).$$

Here $\gamma=10^{-7}\ldots1.0$, $I=H_{k+1}$, $\Delta f_i=M_\alpha\left[f_i\right]-f_i^*$, $\Delta x_p=M_\alpha\left[x_p\right]-x_p^*$, $\chi_{f_i}^2=n_\alpha\dfrac{M_\alpha\left[\left(f_i-M_\alpha\left[f_i\right]\right)^2\right]}{\left(\sigma_{f_i}^*\right)^2}$,

$\chi_{x_p}^2=n_\alpha\dfrac{M_\alpha\left[\left(x_p-M_\alpha\left[x_p\right]\right)^2\right]}{\left(\sigma_{x_p}^*\right)^2}$, $\sigma_{f_i}^0=\dfrac{\sigma_{f_i}}{\sigma_{f_i}^*}$, $\sigma_{x_p}^*=\left\{\dfrac{\sigma_{x_p}}{\sigma_{x_p}^*}\right\}$,

$$\frac{\left|\chi_{f_i}^2-\left(n_\alpha-3\right)\right|}{\sqrt{2\left(n_\alpha-3\right)}}=\frac{n_\alpha}{\sqrt{2\left(n_\alpha-3\right)}}\left|\left(\sigma_{f_i}^*\right)^2-1+\frac{3}{n_\alpha}\right|,$$

$$\frac{\left|\chi_{x_p}^2-\left(n_\alpha-3\right)\right|}{\sqrt{2\left(n_\alpha-3\right)}}=\frac{n_\alpha}{\sqrt{2\left(n_\alpha-3\right)}}\left|\left(\sigma_{x_p}^*\right)^2-1+\frac{3}{n_\alpha}\right|.$$

For the MV-problem solution the computing method based on the memetic algorithm is applied. In this method the evolutional algorithm was realized with the decremental neighborhood method and the randomize path laying method together. The evolutional algorithm has next parameters, which change from epoch to epoch: real coding operator, fitness function and relaxation function.

The proposed method application gives the stable (robust) estimation of desired values under parametric input data uncertainty and lowers the difficulty of the quasisolution synthesis method.

The structure of the model for the choice of a rational individual treatment program for the patients based on IT Data Stream Mining, which implements the “Big Data for Better Outcomes” concept is shown in Fig. 2.

4 EXPERIMENTS

In that work, as an implementation of the concept “Big Data for Better Outcomes” for stratification of patients, the following tasks were solved:

– formation of a subset of controlled variables of state, the values of the quantities of which are check in by measuring instruments;

– robust metamodels: multidimensional diagnostic model;

– informativeness evaluations of the variables of states for different stages of patients’ diseases.

The authors considered the task of assessing the increased survival rate of cancer patients who received chemoradiotherapy.

The multidimensional diagnostic model – model for estimating the survival of patients was built using a data set obtained in the real medical practice of the authors.

The patients with squamous cell carcinomas of the head and cancer (stages III, IVa, IVb) was included in study and did a comprehensive examination, including the collection of anamnestic data, general clinical physical examination, computed tomography of the head, neck and chest organs before the start of radiation therapy and 1 month after its completion, laboratory general clinical and biochemical blood tests. All patient received course of radiation therapy 5 Gy per fraction till 6–7 weeks (32–35

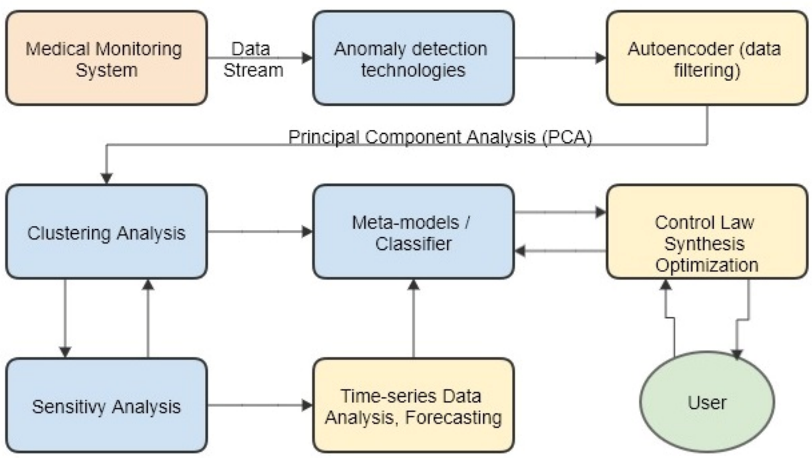

Figure 2 – Context diagram on the implementation of the model of the choice of a rational individual treatment program for the patients, identified on the basis of the forecasted state based on IT Data Stream Mining

DOI 10.15588/1607-3274-2023-1-3

fractions) with chemomodification Cisplatin 100 mg/m^2 at 1, 21, 42 days in SI "Grigoriev Institute for Medical Radiology and Oncology of NAMS of Ukraine". This set contains records of 57 patients on 9 indicators (parameters) that characterize the current state of each patient: age, localization, stage, tumor (T) and lymph (N) nodes status (TNM classification, 8 edition), response (per RECIST 1.1 criteria), tumor marker squamous cell carcinoma's antigen "before" and "after" the course treatment, survival time. All indicators have a numerical type, but some of them are enumerable, for example, the "localization" indicator takes four values from 1 to 4. The objective function for each record is "survival time", the other parameters are attributed to the input data.

First, it is necessary to normalize the data. Since the inputs are different in value and content, the normalization for them will be different. The parameters "localization", "stage", "T" and "response" are normalized according to the formula

$$f_i^0 = \frac{f_i}{f_{\max}},$$

here $f_{\max} = 4$.

The age indicator is normalized as follows:

$$Age_{norm} = \frac{Age_{\max} - Age_{current}}{Age_{\max}},$$

here we take age as the maximum value $Age_{\max} = 85$.

The indicators of tumor markers "before" and "after" are reduced to one parameter of tumor markers as follows:

$$Oncomarker = \ln \frac{marker\_after}{marker\_before}.$$

Let's denote the objective function by *Suv_time*, its values are also pre-normalized by the formula

$$Suv\_time = \frac{time}{target\_time},$$

here $target\_time = 25$.

The results of the patient survival assessment are a functional relationship between the input data and the survival time. Using a robust meta-model based on a radial-basis neural network, survival time estimation values are calculated that correspond to the input data.

The structure of the simplest radial-basis function neural network includes three types of neuron layers. RBF with one hidden layer (K=1). At the input layer RBF has H_0 =7 neurons, at the hidden layer – H_1 = 44 neurons and H_2 =1 network outputs.

In addition, the task of determining the most influential (informative) parameters for assessing patient survival was solved.

Numerical research was carried out with the help of the computer program "Non-linear evaluation methods in multicriteria problem of robust optimal designing and intelligence diagnostics of systems under parametric a priori uncertainty (methodology, methods, techniques and computer systems of support and decision-making implementing them)" (ROD&IDS) developed by the authors [13].

5 RESULTS

To train the neural network and test its operation, the data set of the patient's survival time was divided into two parts: training (47 records, 82%) and test (10 records, 18%).

As a result of training, the relative error was 1.5%.

The relative error on the test sample was 15.4%. The trained neural network can then be used to provide survival estimates for new patients Table 1).

It was found by means of sensitive analysis of variables of the model for each patient from the training sample that on the average the indicators: T, age of the patient, response and stage of the disease are more likely to be informative (Table 2).

Table 2 – The informativeness evaluations of the indicators

	Age	Local.	Stage	T	N	Response	Tumor marker
Suv. time	0.262	0.0025	0.251	0.185	0.0006	0.289	0.009

Therefore, to assess the survival of patients, it is necessary to pay attention to these indicators.

Table 1 – Result of model for test set

Age	Local.	Stage	T	N	Response	Tumor marker	Suv_time (target)	Suv_time (fact)
0.235294	0.75	0.5	1	0.75	0.25	–0.887762	1	1.112889
0.164705	1	1	0	1	0.75	–0.062609	0.28	0.261637
0.364705	1	0.75	2	1	0.5	–0.243848	0.52	0.682176
0.258823	0.25	0.5	1	0.75	0.75	–0.267833	0.44	0.385309
0.352941	0.5	1	0	1	0.75	0.2626409	0.32	0.259509
0.341176	0.75	1	1	1	0.5	–1.659763	0.44	0.387156
0.305882	0.5	0.25	2	1	0.5	–0.212148	0.6	0.591057
0.188235	0.25	0.5	1	0.75	0.25	–0.472130	1	1.026191
0.258823	0.75	0.75	2	1	0.75	–0.819898	0.36	0.512091

DOI 10.15588/1607-3274-2023-1-3

6 DISCUSSION

Thus, the goal of the study was achieved – developed advanced methodology for solving the synthesis problem of the patient's individual treatment program in the medical monitoring systems by means of computational intelligence methods for the medical information analysis and applied information technology realizing them have been developed.

Its application allows:

– reducing the likelihood of adverse states based on the choice of an individual treatment program;

– reducing the probability of incorrectly determining the state of the patients (errors of the third kind in classifying the state of the systems) when monitoring patients;

– obtaining stable effective estimates of unknown values of treatment actions for patients (corresponding to the found state);

– the choice of a rational individual treatment program for the patients, identified on the basis of the forecasted state.

The developed software, which is described below information technology, is being tested for stratification of patients with prostate cancer, squamous cell carcinoma of the head and neck [3–11]. The developed methods for forecasted states are also planned to be verified on the data obtained in various oncological pathologies and to use them in choosing the tactics of treating patients, to forecast treatment complications and assess the patient's curability before and during special treatment.

The material and technical support and personnel base of the Kharkov National Medical University allows collecting data on the course of treatment of cancer patients receiving chemoradiation treatment, using the whole range of non-invasive diagnostic methods, a wide range of tumor markers. Working with patients is regulated by the Bioethics and Deontology Committee and complies with international GCP standards.

As part of the implementation of projects, an innovative strategy is proposed for choosing a rational individual tactics for treating patients based on forecasted states identified using robust stratification methods, which will improve clinical results and prevent complications.

It is also planned to collect data on patients with cancer who have undergone COVID-19 before starting treatment and to assess the effect of the latter on the timing of the start of radiation and chemotherapy, depending on the severity of post Covid's syndrome.

CONCLUSIONS

The scientific novelty obtained results in the fact that an advanced methodology for solving the synthesis problem of a patient's individual treatment program in the medical monitoring systems by means of computational intelligence methods for the medical information analysis and applied information technology realizing them have been developed.

The practical significance of obtained results is that the advanced computational intelligence methods for forecast states can be used in choosing the tactics of treating patients, to forecast treatment complications and assess the patient's curability before and during special treatment.

Prospects for further research are to study application possibilities of the proposed methodology for the choice of a rational individual treatment program for the patients with various chronic diseases, identified on the basis of the forecasted state by means of IT Data Stream Mining.

ACKNOWLEDGEMENTS

The work is supported by the state budget scientific research project of Karazin Kharkiv National University "Information process modeling in the complex and distributed systems" (state registration number 0121U109183) and research projects of Grigoriev Institute for Medical Radiology and Oncology of the National Academy of Medical Sciences of Ukraine "Development of programs for personalized control of the absorbed dose during radiation therapy of genital, head and neck tumors using in vivo dosimetry" (state registration number 0117U001046, 2017 – 2019) and "Optimizing topometrical preparation for radiation therapy of head and neck cancer patients" (state registration number 0119U103013, 2020 – 2022).

REFERENCES

1. Conejar R. J., Kim H.-K. A Medical Decision Support System (DSS) for Ubiquitous Healthcare Diagnosis System, *International Journal of Software Engineering and Its Applications,* 2014, Vol. 8, № 10, pp. 237–244. DOI: http://dx.doi.org/10.14257/ijseia.2014.8.10.22.
2. Wagholikar K. B., Sundararajan V., Deshpande A. W. Modeling Paradigms for Medical Diagnostic Decision Support: A Survey and Future Directions, *Journal of Medical Systems*, 2012, Vol. 36, Issue 5, pp. 3029–3049. DOI: https://doi.org/10.1007/s10916-011-9780-4.
3. Starenkyi V., Goryachaya V., Sokolov O., Ugryumova E. Diagnostic model and information technology of classification states in the differential diagnosis NSCLC (non small cell lung cancer) patients with different methods of radiotherapy and chemotherapy, *Journal of Health Sciences*, 2013, Vol. 3 (8), pp. 7–26. ISSN: 1429-9623.
4. Goriacha V. A., Sokolov O. Yu., Ugryumova K. M., Antonyan I. M., Roshin Y. V., Zelensky A. I., Moshel F. G., Nalbandian T. A. Forecasting of Patients Condition in the \Monitoring Medical Systems on the Example of Prostate Diseases. *Journal of Education, Health and Sport*, 2016, Vol. 6, No. 5, pp. 77–93. DOI: http://dx.doi.org/10.5281/zenodo.51160.
5. Bakumenko N., Strilets V., Ugryumov M. Application of the C-Means Fuzzy Clustering Method for the Patient's State Recognition Problems in the Medicine Monitoring Systems [Electronic resource], *3rd International Conference on Computational Linguistics and Intelligent Systems* (COLINS-2019) proceedings, 2019, Vol. I: Main Conference, 10 p. Access mode: http://ceur-ws.org/Vol-2362/paper20.pdf
6. Strilets V., Bakumenko N., Chernysh S., Ugryumov M., Donets V. Application of artificial neural networks in the problems of the patient's condition diagnosis in medical

DOI 10.15588/1607-3274-2023-1-3

monitoring systems, *Integrated Computer Technologies in Mechanical Engineering – Synergetic Engineering. International Scientific and Technical Conference (ICTM 2019)*, 2020, pp. 173–185. DOI: https://doi.org/10.1007/978-3-030-37618-5_16.

7. Strilets V., Bakumenko N., Donets V., Chernysh S., Ugryumov M., Goncharova T. Machine Learning Methods in Medicine Diagnostics Problem, *16th International Conference on ICT in Education, Research and Industrial Applications. Integration, Harmonization and Knowledge Transfer: proceeding*s, 2020, Vol. II: Workshops, pp. 89–101. Access mode: https://ceur-ws.org/Vol-2732/20200089.pdf.
8. Bakumenko N., Strilets V., Meniailov I., Chernysh S., Ugryumov M., Goncharova T. Synthesis Method of Robust Neural Network Models of Systems and Processes, *Integrated Computer Technologies in Mechanical Engineering*. Springer, Cham, 2020, pp. 3–16. DOI: https://doi.org/10.1007/978-3-030-66717-7_1.
9. Starenkiy V. P., Artiukh S.V., Zelenslyi O. I. and others *Patient's stratification in the medical monitoring systems* : monograph. Kharkiv National Medicine University, 2021, 148 p. (in Ukrainian)
10. Mitryaeva N. A., Starenkiy V. P., Billozor N. V., Grebinyk L. V., Artyukh S. V. Radiosensitization by cyclooxygenase-2 inhibitors in radiation therapy of malignant neoplasms: monograph, *State Institution "Institute of Medical Radiology and Oncology named SP Grigorieva National Academy of Medical Sciences of Ukraine"*, 2021, 136 p.
11. Starenkiy V., Artiukh S., Ugryumov M., Strilets V., Chernysh S., Chumachenko D. A Method for Assessing the Risks of Complications in Chemoradiation Treatment of Squamous Cell Carcinoma of the Head and Neck, *The Open Bioinformatics Journal*, 2021, Vol. 14, pp. 139–143. DOI: 10.2174/18750362021140100138
12. Strilets V., Donets V., Ugryumov M., Zelenskyi R., Goncharova T. Agent-Oriented data clustering for medical monitoring. *Radioelectronic and Computer Systemsthis*, 2022, № 1, P. 103–114. DOI: https://doi.org/10.32620/reks.2022.1.08
13. Ugryumov M. L., Meniaylov Y. S., Chernysh S. V., Ugryumova K. M. (Ukraine) Computer program «*Nonlinear estimation methods in the multicriterion problems of system's robust optimal designing and diagnosing under parametric apriority uncertainty (methodology, methods and computer decision support and making system)*» («ROD&IDS»): Copyright registration certificate № 82875. Copyright and related rights. Official bulletin. Ministry of Economic Development and Trade of Ukraine. 2018, № 51, P. 403.
14. Strilets V. E., Shmatkov S. I., Ugryumov M. L. et al. *Methods of machine learning in the problems of system analysis and decision making*: monograph. Karazin Kharkiv National University, 2020, 195 p. ISBN 978-966-285-627-9. (in Ukrainian)
15. Sidey-Gibbons J. A. M., Sidey-Gibbons Chris J. Machine Learning in medicine: practical introduction, *BMC Medical Research Methodology*, 2019, No. 19. DOI: 10.1186/s12874-019-0681-4.
16. Varoquaux G., Cheplygina V. Machine Learning for medical imaging: methodological failures and recommendations for the future, *Digit. Med.*, 2022, No. 5 (48). DOI: https://doi.org/10.1038/s41746-022-00592-y.
17. Antoniou T., Mamdani M. Evaluation of machine learning solutions in medicine, *CMAJ: Canadian medical association Journal*, 2021, No. 193 (36), E1425–E1429. DOI: 10.1503/cmaj.210036.
18. Seyhan A. A., Carini C. Are innovation and new technologies in precision medicine paving a new era in patients centric care? *Journal of Translational Medicine*, 2019, No. 17, article number 114. DOI: https://doi.org/10.1186/s12967-019-1864-9
19. Nayyar Anand, Gadhavi Lata, Zaman Noor. Chapter 2 – Machine learning in healthcare: review, opportunities and challenges, *Machine Learning and the Internet Medical Things in Healthcare*, 2021, P. 23–45. DOI: https://doi.org/10.1016/B978-0-12-821229-5.00011-2.
20. Trujillano J., Badia M., Serviá L. et al. Stratification of the severity of critically ill patients with classification trees, *BMC Medical Research Methodology*, 2009, Vol. 9, No. 83. DOI: https://doi.org/10.1186/1471-2288-9-83.
21. Amigo J. M., Small M. Mathematical methods in medicine: Neuroscience, cardiology and pathology, *Philosophical Transactions of The Royal Society A Mathematical Physical and Engineering Sciences*, 2017, Vol. 375 (2096): 20170016. DOI: 10.1098/rsta.2017.0016.
22. Azar A., El-Metwally S. Decision tree classifiers for automated medical diagnosis, *Neural Computing and Applications*, 2013, Vol. 23, pp. 2387–2403. DOI: 10.1007/s00521-012-1196-6.
23. Bahari N. I. S., Ahmad A., Aboobaider B. M. Application of support vector machine for classification of multispectral data, *OP Conference Series: Earth and Environmental Science*, 2014, Vol. 20 (1): 012038. DOI: 10.1088/1755-1315/20/1/012038.
24. Friedman J., Hastie T., Tibshirani R. Additive logistic regression: a statistical view of boosting, *Ann. Statist.*, 2000, Vol. 28, № 2, pp. 337–407. DOI: 10.1214/aos/1016218223.
25. Vivek Verma. Application of Bayesian Analysis in Medical Diagnosis, *Journal of the Practice of Cardiovascular Science*, 2019, Vol. 5, P. 136–141. DOI: 10.4103/jpcs.jpcs_51_19.
26. Alfonso Perez G., Caballero Villarraso J. Alzheimer Identification through DNA Methylation and Artificial Intelligence Techniques, *Mathematics*, 2021, No. 9, 2482. DOI: https://doi.org/10.3390/math9192482.

Received 31.10.2022.
Accepted 22.12.2022.

УДК 519.876.5:681.518.2

ОБЧИСЛЮВАЛЬНІ ІНТЕЛЕКТУАЛЬНІ МЕТОДИ СТРАТИФІКАЦІЇ ПАЦІЄНТІВ У СИСТЕМАХ МЕДИЧНОГО МОНІТОРИНГУ

Бакуменко Н. С. – канд. техн. наук, доцент, доцент кафедри теоретичної та прикладної системотехніки Харківського національного університету імені В.Н. Каразіна, Харків, Україна.

Стрілець В. Є. – канд. техн. наук, доцент кафедри теоретичної та прикладної системотехніки Харківського національного університету імені В. Н. Каразіна, Харків, Україна.

DOI 10.15588/1607-3274-2023-1-3

Угрюмов М. Л. – д-р техн. наук, професор, професор кафедри теоретичної та прикладної системотехніки Харківського національного університету імені В. Н. Каразіна, Харків, Україна.

Зеленський Р. О. – асистент кафедри урології та дитячої урології Харківської медичної академії післядипломної освіти, Харків, Україна.

Угрюмова К. М. – канд. техн. наук, доцент кафедри математичного моделювання та штучного інтелекту Національного аерокосмічного університету ім. М.Є. Жуковського «Харківський авіаційний інститут», Харків, Україна.

Старенький В. П. – д-р мед. наук, професор, завідувач відділення радіаційної онкології, ДУ «Інститут медичної радіології та онкології ім. С.П. Григор’єва» Національної академії медичних наук України, Харків, Україна.

Артюх С. В. – канд. мед. наук, старший науковий співробітник групи променевої терапії відділу радіології ДУ «Інститут медичної радіології та онкології ім. С.П. Григор’єва» НАМН України, асистент кафедри радіології та радіаційної медицини Харківського національного медичного університету, Харків, Україна.

Насонова А. М. – канд. мед. наук, клінічний онколог, завідувачка відділенням клінічної онкології та гематології ДУ «Інститут медичної радіології та онкології ім. С.П. Григор’єва» НАМН України, Харків, Україна.

АНОТАЦІЯ

Актуальність. У сучасній медичній практиці все більше впроваджується автоматизація й інформаційні технології для діагностування захворювань, моніторингу стану пацієнта, визначення програми лікування тощо. Тому розробка нових і удосконалення існуючих методів стратифікації пацієнтів у системах медичного моніторингу є своєчасною і необхідною.

Метод. Розроблено методологію, яка включає такі методи обчислювального інтелекту для стратифікації пацієнтів у системах медичного моніторингу, як:

1) метод кластерного аналізу на основі агентного підходу – визначення можливої кількості станів пацієнтів з використанням контрольованих змінних станів;

2) метод побудови робастних метамоделей за допомогою штучних нейронних мереж при апріорній невизначеності даних (відома лише точність вимірювань) за даними моніторингу стану пацієнтів: а) багатовимірна логістична регресійна модель у вигляді аналітичних залежностей апостеріорних ймовірностей різних станів пацієнтів від контрольованих змінних станів; б) багатовимірна діагностична модель у вигляді аналітичних залежностей цільових функцій (критеріїв якості стану хворого) від контрольованих змінних станів;

3) метод оцінки інформативності контрольованих змінних станів при невизначеності апріорних даних;

4) метод побудови робастних багатовимірних моделей контролю стану пацієнтів при апріорній невизначеності даних у даних моніторингу у вигляді аналітичних залежностей, що прогнозуються за виміряними значеннями контрольованих змінних станів у процесі моніторингу;

5) метод зменшення розмірності простору контрольованих змінних станів на основі аналізу інформативності змінних робастних багатовимірних моделей управління станом пацієнтів (аналіз чутливості);

6) метод визначення станів пацієнтів на основі вирішення задачі класифікації за значеннями контрольних та прогнозованих контрольованих змінних стану з використанням імовірнісних нейронних мереж;

7) метод синтезу раціональної індивідуальної програми лікування хворих у системі медичного моніторингу для стану, виявленого на основі прогнозу.

У роботі запропонована структура моделі вибору раціональної індивідуальної програми лікування пацієнтів на основі IT Data Stream Mining, яка реалізує концепцію « Big Data for Better Outcomes».

Результати. Розроблені передові методи обчислювального інтелекту для прогнозування станів використовувалися при виборі тактики лікування пацієнтів, прогнозуванні ускладнень лікування та оцінці виліковності пацієнта до та під час спеціального лікування.

Висновки. Представлено досвід впровадження концепції «Big Data for Better Outcomes» для вирішення проблеми розробки передових методологій нових стратегій стратифікації пацієнтів. Розроблено передову методологію, методи обчислювального інтелекту для стратифікації пацієнтів у системах медичного моніторингу та прикладну інформаційну технологію, що їх реалізує. Розроблені передові методи прогнозування станів можуть бути використані при виборі тактики лікування хворих, прогнозуванні ускладнень лікування та оцінці виліковності хворого до та під час спеціального лікування.

КЛЮЧОВІ СЛОВА: інформаційна технологія потокового аналізу даних, системи медичного моніторингу, методи машинного навчання, математичні моделі і методи стратифікації пацієнтів.

УДК 519.876.5:681.518.2

ВЫЧИСЛИТЕЛЬНЫЕ ИНТЕЛЛЕКТУАЛЬНЫЕ МЕТОДЫ СТРАТИФИКАЦИИ ПАЦИЕНТОВ В СИСТЕМАХ МЕДИЦИНСКОГО МОНИТОРИНГА

Бакуменко Н. С. – канд. техн. наук, доцент, доцент кафедры теоретической и прикладной системотехники Харьковского национального университета имени В. Н. Каразина, Харьков, Украина.

Стрелец В. Е. – канд. техн. наук, доцент кафедры теоретической и прикладной системотехники Харьковского национального университета имени В. Н. Каразина, Харьков, Украина.

Угрюмов М. Л. – д-р техн. наук, профессор, профессор кафедры теоретической и прикладной системотехники Харьковского национального университета имени В. Н. Каразина, Харьков, Украина.

Зеленский Р. А. – ассистент кафедры урологии и детской урологии Харьковской медицинской академии последипломного образования, Харьков, Украина.

DOI 10.15588/1607-3274-2023-1-3

Угрюмова Е. М. – канд. техн. наук, доцент кафедры математического моделирования и искусственного интеллекта Национального аэрокосмического университета им. Н.Е. Жуковского «Харьковский авиационный институт», Харьков, Украина.

Старенький В. П. – д-р мед. наук, профессор, заведующий отделением радиационной онкологии, ГО «Институт медицинской радиологии и онкологии им. П. С. Григорьева» Национальной академии медицинских наук Украины, Харьков, Украина.

Артюх С. В. – канд. мед. наук, старший научный сотрудник группы лучевой терапии отдела радиологии ГО «Институт медицинской радиологии и онкологии им. П. С. Григорьева» Национальной академии медицинских наук Украины, ассистент кафедры радиологии и радиационной медицины, Харьковский национальный медицинский университет, Харьков, Украина.

Насонова А. Н. – канд. мед. наук, клинический онколог, заведующая отделением клинической онкологии и гематологии ГО «Институт медицинской радиологии и онкологии им. П. С. Григорьева» Национальной академии медицинских наук Украины, Харьков, Украина.

АННОТАЦИЯ

Актуальность. В современной медицинской практике все больше внедряются автоматизированные и информационные технологии для диагностики заболеваний, мониторинга состояния пациентов, выбора программы лечения и др. В связи с этим разработка новых и усовершенствование существующих методов вычислительного интеллекта для стратификации пациентов в системах медицинского мониторинга является своевременной и необходимой задачей.

Метод. Разработана методология, включающая следующие методы вычислительного интеллекта для стратификации пациентов в системах медицинского мониторинга:

1) метод кластерного анализа на основе агентного подхода – определение возможного количества состояний пациентов с использованием контролируемых переменных состояния;

2) метод построения робастных метамоделей с помощью искусственных нейронных сетей при априорной неопределенности данных мониторинга (известна только точность измерений) в данных мониторинга: а) многомерная логистическая регрессионная модель в виде аналитических зависимостей апостериорных вероятностей различных состояний пациентов от контрольных и контролируемых переменных состояния; б) многомерная диагностическая модель в виде аналитических зависимостей целевых функций (критериев качества состояния больного) от контрольных и контролируемых переменных состояния.

3) метод оценки информативности контролируемых переменных состояния при неопределенности априорных данных;

4) метод построения робастных многомерных моделей контроля состояния пациентов при априорной неопределенности данных мониторинга в виде аналитических зависимостей, прогнозируемых по измеренным значениям контрольных и контролируемых переменных состояния в процессе мониторинга;

5) метод уменьшения размерности пространства контролируемых переменных состояния на основе анализа информативности переменных робастных многомерных моделей управления состоянием пациентов;

6) метод определения состояний пациентов на основе решения задачи классификации по значениям контрольных и прогнозируемых контролируемых переменных состояния с использованием вероятностных нейронных сетей;

7) метод синтеза рациональной индивидуальной программы лечения больных в системе медицинского мониторинга, для состояния, выявленного на основе прогноза.

В работе предложена структура модели выбора рациональной индивидуальной программы лечения пациентов на основе IT Data Stream Mining, которая реализует концепцию «Big Data for Better Outcomes».

Результаты. Разработанные усовершенствованные методы вычислительного интеллекта для прогноза состояний использовались при выборе тактики лечения пациентов, для прогнозирования осложнений лечения и оценки излечимости пациента до и во время проведения специального лечения.

Выводы. Представлен опыт реализации концепции «Большие данные для лучших результатов» для решения задачи разработки передовых методологий для новых стратегий стратификации пациентов. Разработаны передовая методология, методы вычислительного интеллекта для стратификации пациентов в системах медицинского мониторинга и реализующие их прикладные информационные технологии. Разработанные усовершенствованные методы прогноза состояний могут быть использованы при выборе тактики лечения больных, для прогнозирования осложнений лечения и оценки излечимости больного до и во время проведения специального лечения.

КЛЮЧЕВЫЕ СЛОВА: информационная технология потокового анализа данных, системы медицинского мониторинга, методы машинного обучения, математические модели и методы стратификации пациентов.

ЛІТЕРАТУРА

1. Conejar R. J. A Medical Decision Support System (DSS) for Ubiquitous Healthcare Diagnosis System / R. J. Conejar, H.-K. Kim // International Journal of Software Engineering and Its Applications. – 2014. – Vol. 8, № 10. – P. 237–244. DOI: http://dx.doi.org/10.14257/ijseia.2014.8.10.22.
2. Wagholikar K. B. Modeling Paradigms for Medical Diagnostic Decision Support: A Survey and Future Directions / K. B. Wagholikar, V. Sundararajan, A. W. Deshpande // Journal of Medical Systems. – 2012. – Vol. 36, Issue 5. – P. 3029–3049. DOI: https://doi.org/10.1007/s10916-011-9780-4.
3. Diagnostic model and information technology of classification states in the differential diagnosis NSCLC (non small cell lung cancer) patients with different methods of radiotherapy and chemotherapy / [V. Starenkyi, V. Goryachaya, O. Sokolov, E. Ugryumova] // Journal of Health Sciences. – 2013. – Vol. 3 (8). – P. 7–26. ISSN: 1429-9623.
4. Forecasting of Patients Condition in the Monitoring Medical Systems on the Example of Prostate Diseases / [V. A. Goriacha, O. Yu. Sokolov, K. M. Ugryumova et al.] // Journal of

DOI 10.15588/1607-3274-2023-1-3

Education, Health and Sport. – 2016. – Vol. 6, № 5. – P. 77–93. DOI: http://dx.doi.org/10.5281/zenodo.51160.
5. Bakumenko N. Application of the C-Means Fuzzy Clustering Method for the Patient's State Recognition Problems in the Medicine Monitoring Systems [Electronic resource] / N. Bakumenko, V. Strilets, M. Ugryumov // 3rd International Conference on Computational Linguistics and Intelligent Systems (COLINS-2019), Kharkiv, Ukraine, April 18–19, 2019: proceedings. – 2019. – Vol. I: Main Conference. – 10 p. Access mode: http://ceur-ws.org/Vol-2362/paper20.pdf
6. Application of artificial neural networks in the problems of the patient's condition diagnosis in medical monitoring systems / [V. Strilets, N. Bakumenko, S. Chernysh et al.] // Integrated Computer Technologies in Mechanical Engineering – Synergetic Engineering. International Scientific and Technical Conference (ICTM 2019). Kharkov, Ukraine, November, 28–30, 2019. – Springer Link: 2020. – P. 173–185. DOI: https://doi.org/10.1007/978-3-030-37618-5_16.
7. Machine Learning Methods in Medicine Diagnostics Problem / [V. Strilets, N. Bakumenko, V. Donets et al.] // 16th International Conference on ICT in Education, Research and Industrial Applications. Integration, Harmonization and Knowledge Transfer, Kharkiv, Ukraine, 6–10 October 2020: proceedings. – Volume II: Workshops, 2020. – P. 89–101. Access mode: https://ceur-ws.org/Vol-2732/20200089.pdf.
8. Bakumenko N. Synthesis Method of Robust Neural Network Models of Systems and Processes / [N. Bakumenko, V. Strilets, I. Meniailov et al.] // Integrated Computer Technologies in Mechanical Engineering – 2020 / M. Nechyporuk, V. Pavlikov, D. Kritskiy (eds) – Springer, Cham. – P. 3–16. DOI: https://doi.org/10.1007/978-3-030-66717-7_1.
9. Стратифікація пацієнтів у системах медичного моніторингу : монографія / [В. П. Старенький, С. В. Артюх, О. І. Зеленський та ін.]; под ред. О. Г. Авруніна, Є. Б. Радзишевської. – Харків : ХНМУ, 2021. – 148 с.
10. Radiosensitization by cyclooxygenase-2 inhibitors in radiation therapy of malignant neoplasms: monograph / [N. A. Mitryaeva, V. P. Starenkiy, N. V. Billozor et al.]. – Kharkiv : State Institution "Institute of Medical Radiology and Oncology named SP Grigorieva National Academy of Medical Sciences of Ukraine", 2021. – 136 p.
11. A Method for Assessing the Risks of Complications in Chemoradiation Treatment of Squamous Cell Carcinoma of the Head and Neck / [V. Starenkiy, S. Artiukh, M. Ugryumov et al.] // The Open Bioinformatics Journal. – 2021. – Vol. 14. – P. 139–143. DOI: 10.2174/18750362021140100138.
12. Agent-Oriented data clustering for medical monitoring / [V. Strilets, V. Donets, M. Ugryumov et al.] // Radioelectronic and Computer Systemsthis. 2022. – № 1. – P. 103–114. DOI: https://doi.org/10.32620/reks.2022.1.08
13. Комп'ютерна програма «Методи нелінійного оцінювання в багатокритеріальних задачах робастного оптимального проектування та діагностування систем в умовах параметричної апріорної невизначеності (методологія, методи, методики і реалізують їх комп'ютерні системи підтримки і прийняття рішень)» («ROD&IDS»): Свідоцтво про реєстрацію авторського права на твір № 82875 / М. Л. Угрюмов, Є. С. Меняйлов, С.В.Черниш, К. М. Угрюмова (Україна). – Авторське право і суміжні права. Офіційний бюлетень. – К.: Міністерство економічного розвитку і торгівлі України, 2018. – № 51. – С. 403.
14. Методи машинного навчання у задачах системного аналізу і прийняття рішень: монографія / В. Є. Стрілець, С. І. Шматков, М. Л. Угрюмов та ін.]. – Харків : Харківський національний університет імені В. Н. Каразіна, 2020. – 195 с. – ISBN 978-966-285-627-9.
15. Sidey-Gibbons J. A. M. Machine Learning in medicine: practical introduction / Jenny A.M. Sidey-Gibbons, Chris J Sidey-Gibbons // BMC Medical Research Methodology. – 2019. – No. 19. DOI: 10.1186/s12874-019-0681-4.
16. Varoquaux G. Machine Learning for medical imaging: methodological failures and recommendations for the future / G. Varoquaux, V. Cheplygina // Digit. Med. – 2022. – No. 5(48). DOI: https://doi.org/10.1038/s41746-022-00592-y.
17. Antoniou T. Evaluation of machine learning solutions in medicine / T. Antoniou, M. Mamdani // CMAJ: Canadian medical association Journal. – 2021. – No. 193(36). – E1425-E1429. DOI: 10.1503/cmaj.210036.
18. Seyhan A. A. Are innovation and new technologies in precision medicine paving a new era in patients centric care? / A. A. Seyhan, C. Carini // Journal of Translational Medicine. – 2019. – № 17. – Article 114. DOI: https://doi.org/10.1186/s12967-019-1864-9
19. Nayyar Anand. Chapter 2 – Machine learning in healthcare: review, opportunities and challenges / Nayyar Anand, Gadhavi Lata, Zaman Noor // Machine Learning and the Internet Medical Things in Healthcare. – 2021. – P. 23–45. DOI: https://doi.org/10.1016/B978-0-12-821229-5.00011-2.
20. Stratification of the severity of critically ill patients with classification trees / [J. Trujillano, M. Badia, L. Serviá et al.] // BMC Medical Research Methodology. – 2009. – Vol. 9, No. 83. DOI: https://doi.org/10.1186/1471-2288-9-83.
21. Amigo J. M. Mathematical methods in medicine: Neuroscience, cardiology and pathology / J. M. Amigo, M. Small // Philosophical Transactions of The Royal Society A Mathematical Physical and Engineering Sciences. – 2017. – vol. 375 (2096):20170016. DOI: 10.1098/rsta.2017.0016.
22. Azar A. Decision tree classifiers for automated medical diagnosis / A. Azar, S. El-Metwally // Neural Computing and Applications. – 2013. – Vol. 23. – P. 2387–2403. DOI: 10.1007/s00521-012-1196-6.
23. Bahari N. I. S. Application of support vector machine for classification of multispectral data / N. I. S. Bahari, A. Ahmad, B. M. Aboobaider // OP Conference Series: Earth and Environmental Science. – 2014. – Vol. 20 (1): 012038. DOI: 10.1088/1755-1315/20/1/012038
24. Friedman J. Additive logistic regression: a statistical view of boosting / J. Friedman, T. Hastie, R. Tibshirani // Ann. Statist. – 2000. – Vol. 28, № 2. – P. 337–407. DOI: 10.1214/aos/1016218223.
25. Vivek Verma. Application of Bayesian Analysis in Medical Diagnosis / Vivek Verma // Journal of the Practice of Cardiovascular Science. – 2019. – Vol. 5. – P. 136–141. DOI: 10.4103/jpcs.jpcs_51_19.
26. Alfonso Perez G. Alzheimer Identification through DNA Methylation and Artificial Intelligence Techniques / G. Alfonso Perez, J. Caballero Villarraso // Mathematics. – 2021. – № 9. – 2482. DOI: https://doi.org/10.3390/math9192482.

DOI 10.15588/1607-3274-2023-1-3