Journal of Health Sciences. 2013; 3 (8): 7-26
The journal has had 4 points in Ministry of Science and Higher Education of Poland parametric evaluation. Part B item 683.

Conflict of interest: None declared. Received: 25.04.2013. Revised: 25.07.2013. Accepted: 08.08.2013.

UDC: 681.518.2:519.816

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

Victor Starenkyi[1], Veronica Goryachaya[2], Oleksandr Sokolov[3], Ekatherina Ugryumova[2]

[1]Department of Radiotherapy, Institute of Medical Radiology named after S.P. Grigoriev, Pushkinskaya Street 82, 61024, Kharkiv, Ukraine
E-mail: starenkiyvp@rambler.ru
[2]Department of Informatics, National Aerospace University Chkalov Street 17, 61070, Kharkiv, Ukraine
E-mail: splaijna@gmail.com
E-mail: ugryumov.m@mail.ru
[3]Department of Informatics, Faculty of Physics, Astronomy and Informatics, Nicolaus Copernicus University, Grudzianska 5, 87-100, Torun, Poland
E-mail: osokolov@islumk.pl

Abstract: The method for solving the classification condition of elements complex systems has been improved. The system model of the diagnosing medical process and biological elements of the system (patients) has been proposed. Based on the information content analysis and aggregating state variables made the reduction of the dimension variables of the status patients. Computer decision support system for biomedical diagnose condition of patients has been developed.

Keywords: regression analysis, reduction of dimension, classification states.

1. Statement of the problem and its relevance

Epidemiological studies, which have been conducted in different countries, indicate a gradual increase in the frequency incidence of lung cancer, but in this investigated sampling are increases of severe cases (IIIA – IIIB stages).

Treatment planning and statistical of obtained data analysis are necessary part of scientific activity. Statistical methods make it possible not only to evaluate new methods of diagnosis and treatment, but also to supplement the information content of traditional indicators, which are critical to making a decision regarding the choice of tactics of radiation and chemoradiation therapy.

The development of defects in engineering or biomedical systems is a complex dynamic process. Subject area experts cannot always predict how quickly they will be develop. It is not always possible to reach a consensus at what stage of development are located the defects and, as a consequence of which methods of eliminating defects needs to apply. Monitoring and forecasting dynamic process of the system work, help to experts and/or patients (if biomedical systems are considered) to make decisions that lead to better values criteria quality of the work system or the survival and quality of life for patients.

As an example of a complex system, we consider medico-biological system, which consists of the following elements: doctor, patients and subsystem diagnosis of patients. The system model process of diagnosing biomedical system is shown in Figure 1, where Database – the database of patients with measurable parameters; S – the adder of an inputs; Controller – the governing body (it's a doctor, that develops a script treatment of patients); u – the control variables; f – external influences (disturbance); Object of control – the object of management (patients); X – the state variables patients; Y – the quality criteria of the states patients; Anamnez – data collection anamnesis; Laboratory diagnosis – data of laboratory diagnosis; Visual diagnosis – data visual diagnostics; The diagnosis – establishment of the diagnosis.

The beginning and the end stages of the life cycle of the treatment patients we defined as the set of final states of the patients. The number of states have taken to address is set by the expert in the subject area based on the results of the cluster analysis. It has accepted as a working hypothesis, the hypothesis of local equilibrium, according to which the patient is uniquely determined, with the fundamental system his state variables.

We assume that the critical on the set of final states is a resistant state in which, due to the progressive development of defects in the functional parts, the patient in the treatment process becomes unmanageable.

The objective problem is that there is no structured allowing rule transition for the patient in question in a resistant state.

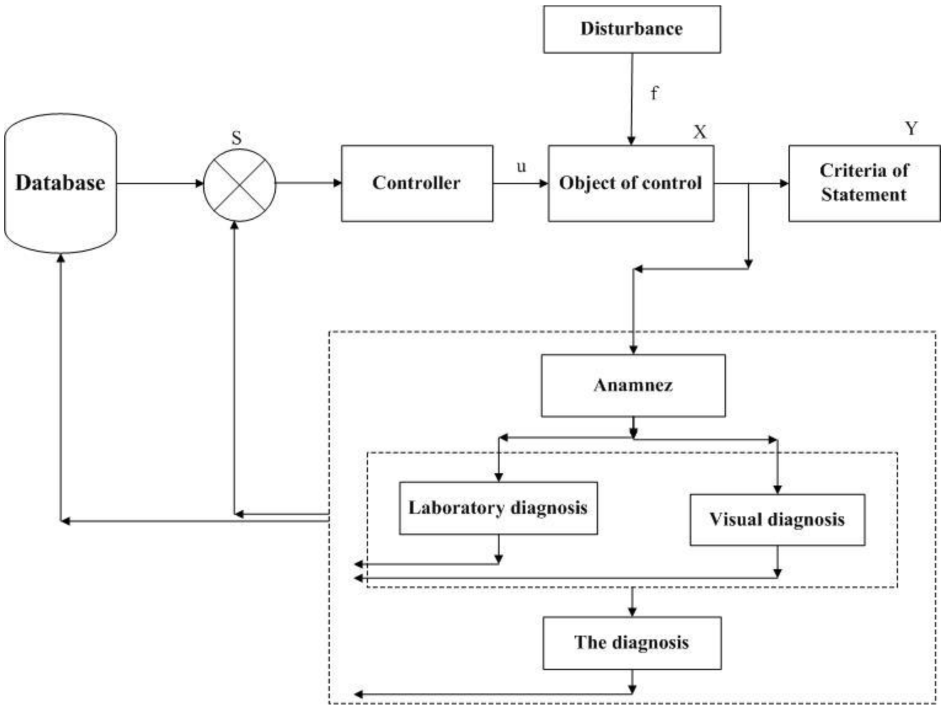

Fig. 1. System model process of diagnosing medical and biological systems.

Great attention is given consideration of problems in the theory and practice of diagnosing the state of complex systems by both scholars in Ukraine and abroad. A lot of papers has been published to the present time, produced describing the methods for solving problems of diagnosing technical and medical-biological systems (see, for example [1-10]). In the papers [1-3] are sufficiently described the general provisions of the trainees artificial neural network (ANN) theory, which are widely used for the construction of formal mathematical models (diagnostic models) in the form of regression equations. The prospects of application and development classification methods of state (pattern recognition) has been discussed in the papers [4-10], such as: the formation of informative state variables; classification based on the training set; accounting dynamics of the state variables of control objects.

It should be noted that in the majority of works devoted to the diagnostic challenges and prediction based on the ANN, there are no variables information content analysis of diagnostic models.

Analysis of existing sources literature shows that in the development of the problem-solving diagnosing methods, such as building a diagnostic model, classification and prediction of the state, causes a number of mathematical problems such as:

- large dimension of problems;
- the state variables are correlated variables (multicollinearity);
- low information content of state variables;
- there is uncertainty in the formation and structuring of decision rules status classification systems elements.

To ensure the reliability of the information obtained through regression models, these models should have the property of stability in the conditions of a priori uncertainty of the input data, as well as sufficient from a practical point of view of accuracy.

Information support, which has been developed for today does not allow with a sufficiently high level of confidence to solve the classification problem for the elements of complex systems.

Thus, there is a need to improvement models, mathematical methods and information technology tools processing of statistical data for the states analysis of elements dynamic systems.

The purpose of this study is to analyze the statistical relationship between the variables that determine the state of the elements biomedical systems (patients); Identification affiliations patient to a particular class based on its conditions sensed variables.

During the research process were solving the following scientific objectives: the construction of the diagnostic model; reduction of dimension vector input data; aggregation of state variables into complexes, which pairwise would have been poorly correlated variables (orthogonalization); classification state of patients.

2. Staging and method of solving the problem of constructing the diagnostic model

Based on a systematic analysis of process diagnosing was revealed hierarchy of stages diagnosing: laboratory diagnosis (biochemical blood tests, etc.), visual diagnostics (ultrasound, MRI, etc.) and identified corresponding to each step registered variables of the patients. The initial dimension of the space state variables was equal to 11.

Further, experimental sample of sensed variables that characterize the state of the observed patients was formed. Hereinafter the sample was divided into four classes according to the TNM classification:

- healthy patients – 81 people;
- “The first and the second stages” (It’s a small or large tumor size in the parenchyma of the organ, but it does not affect the capsule organ or invades the muscle layer of the hollow body, an isolated metastasis in the lymph nodes) – 175 patients;

- “The third stage” (It’s a tumor of considerable size and affects all tissue parenchymal organ or germinates through the all the layers of a hollow organ, numerous metastases which are usually linked with each other and with the tissues in the lymph nodes) – 194 patients;
- “The fourth stage” (the tumor extends beyond the organ then germinates into the surrounding organs and tissues, usually it comes to the regional lymph nodes that are inaccessible (or difficult to access) to remove) – 178 patients.

The data are summarized in Excel Table, as shown in Table I. As an indication of the classification by dividing the total sample into classes was chosen level of disease progression.

Correlation analysis had been performed to identify the statistical dependence between random variables that determine the condition of the patients, through the point estimate of the correlation coefficients for each of the classes. In what follows a variable HCE_NL was removed from the analysis because of the low reliability (it means that the majority patients had the missed data of analysis or values hadn't been accurate).

For selected variables has been analyzed trends of the mean values of the state variables in classes (see Excel table in Table I). The trend of changes detected in the following variables with increasing levels of disease progression: bombesin, PEA, the Karnofsky scale. The Karnofsky scale (KS) had been chosen as a criterion quality of the patients.

Table I. Baseline patients’ data for the formation of the variables that determine their state

1	2	3	4	5	6	7	8	9	10	11
DR	POL	M_STR	ORG_OM	T	N	B_NL	REA_NL	DL	FORM	KP
63,00	1,00	0,00	0,00	0,00	0,00	42,30	8,71	0,00	0,00	96
51,00	1,00	0,00	0,00	0,00	0,00	39,80	8,20	0,00	0,00	98
49,00	1,00	0,00	0,00	0,00	0,00	41,10	8,90	0,00	0,00	96
56,00	1,00	0,00	0,00	0,00	0,00	42,30	8,71	0,00	0,00	92
…										

51,00	1,00	1,00	2,00	2,00	1,00	157,00	66,00	3,00	2,00	90,00
49,00	1,00	1,00	4,00	2,00	1,00	166,00	36,50	1,00	1,00	85,00
56,00	1,00	1,00	1,00	2,00	1,00	198,00	74,30	1,00	1,00	90,00
72,00	1,00	1,00	5,00	2,00	1,00	186,30	37,00	3,00	2,00	90,00
54,00	1,00	1,00	4,00	2,00	1,00	161,00	74,50	3,00	1,00	90,00
…										
63,92	1,00	1,00	6,00	1,00	2,00	234,80	73,20	1,00	1,00	80,00
58,51	1,00	1,00	1,00	3,00	2,00	183,00	86,40	2,00	1,00	70,00
44,67	1,00	1,00	2,00	3,00	2,00	191,60	91,50	2,00	2,00	70,00
70,49	1,00	2,00	5,00	3,00	2,00	185,80	85,60	2,00	1,00	90,00
57,99	1,00	2,00	6,00	3,00	1,00	192,50	87,30	2,00	1,00	80,00
…										
67,13	1,00	2,00	6,00	4,00	2,00	241,00	131,00	4,00	1,00	76,00
58,46	1,00	2,00	4,00	4,00	2,00	260,00	109,00	1,00	1,00	80,00
43,20	1,00	2,00	6,00	4,00	1,00	244,00	129,00	2,00	2,00	80,00
63,14	1,00	1,00	5,00	4,00	2,00	266,00	86,00	1,00	1,00	70,00
56,83	1,00	1,00	5,00	4,00	2,00	243,00	131,00	1,00	2,00	70,00

		Group number			
Name of variables	Meaninig	0	1	2	3
DR	Age	61,6543	61,64486	61,7889	61,44246
POL	Gender	1,135802	1,068571	1,221649	1,095506
M_STR	Metastases	0	1,4	1,396907	1,370787
ORG_OM	Metastatic organ	0	4,622857	4,438144	4,662921
T	Stage T	0	2,434286	2,427835	3,988764
N	Stage N	0	0,577143	1,701031	1,123596
B_NL	Bombesin	40,94815	183,2109	208,0531	243,4888
REA_NL	REA	7,884321	53,86686	78,61546	97,10899
DL	Lung lobe	0	1,485714	1,474227	1,477528
FORM	Type of the tumor	0	1,4	1,438144	1,404494
KP	Karnofsky scale	95,17607	81,71429	80,47423	76,4382

Fig 2. The average values of the state variables.

Suppose there is a multi-dimensional matrix of states $X = \{x_{i,j}\}$ $(i = 1..I, j = 1..J)$, where I – the number of observed patients in the sample, J – the number of measured state variables. Traditionally, the rows of this matrix are called precedents. We perform the centering and standardization of data:

$$x_{ij}^{\circ} = \left(x_{ij} - \langle X_j \rangle\right) / \sigma_j \, , \qquad (1)$$

where $\langle X_j \rangle$ – average of the j's state variable; σ_j – its standard deviation.

We formulate the statement of the problem constructing a diagnostic model. There is given a vector function through a set of training pairs $\left(\vec{X}^{(0)}, \vec{d}\right)_p$, $p = 1..P$, where $\vec{X}^{(0)}, \vec{d}$ – input vectors with dimension H_0, and output vector with dimension H_{K+1}, respectively. This sample must be approximated. The result of solving the problem should be a mathematical mechanism, as a result of which work it was possible to get any value of the vector function $\vec{Y}^{(K+1)}\left(\vec{X}^{(0)}\right)$, which is represented by a given training set based on a specified input vector, in the range which is limited by the input data.

The correlation and regression analysis methods [10-12] were used in order to identify linkages and regularities in development researched process.

Correlation analysis. In order to detect statistical dependence between random variables that determine the condition of the patients were obtained point estimates coefficients of pair correlations for each of the classes. The quality criterion has been chosen scale condition of the patients. Based on the analysis of pair correlation coefficients were identified variables for each of the classes for which the pair correlation coefficients exceed the selected level of significance (> 0.15) [11]. Thus, the state variables have been identified, the impact of which on a scale condition revealed the determining factor.

Regression analysis. Based on normalized variables with using a generalized least squares (OLS) was obtained by linear multiple regression equation in the form: $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_J X_J + \varepsilon$, where β_j ($j = 1..J$) – coefficients of linear multiple regression (they are presented in the Table II).

Table II. *Coefficients of the equation multiple linear regression*

beta1	-0,038339
beta2	-0,022420
beta3	-0,138732
beta4	-0,104550
beta5	-0,494796

beta6	-0,109025
beta7	0,183213
beta8	0,154101
beta9	-0,197986
beta10	-0,070018

Evaluation of information content variables of the diagnostic model in the form of coefficients significance of the contribution of the state variables in the values of the quality criterion – $\varphi_{\beta_j} = \frac{\beta_j^2}{\sum_{j=0}^{J} \beta_j^2}$, was carried out using the parameters found of the regression equation, according to the criterion of the conditional entropy [2]: $I(Y|X_j, X_j) = \frac{1}{2}\log_2\left(\frac{\sigma^2_{Y|X_j}}{\sigma^2_{X_j}}\right) = \frac{1}{2}\log_2 \beta_j^2$. In the result analysis of variables information content had been revealed that the dimension space of the state variables can be reduced to 5 to 10 variables. The coefficients of the importance of the contribution of the state variables to the values KS are shown in the Table. III.

Table III. *The coefficients the contribution importance of the state variables*

T	0,6273
DL	0,1004
B_NL	0,0860
REA_NL	0,0608
M_STR	0,0493
N	0,0305
ORG_OM	0,0280
FORM	0,0126
DR	0,0038
POL	0,0013

The coefficient of determination has equaled 0,384548; the multiple correlation coefficient is equal 0,62012.

Thus, based on conducted analysis has been revealed 5 of 10 informative variables that determine the condition of the patients in classes according to the chosen criterion quality of the scale condition of the patients.

3. The statement and method of solving classification problem state elements of dynamic systems

We formulate the statement of the classification problem. Let the $\vec{X}^*$ be is a vector of variables describing the state of the precedents and *M* – it is the set of numbers classes (of scenarios). It is known the count of possible failure scenarios system in the whole and for each scenario (class) had been formed subset of observed state variables (symptoms). According to the values of the vector $\vec{X}^*$ projections the precedent is assigned to one of the possible sets R_m, где $m = 0..M-1$. Is needed to find a *m*-th scenario, where precedent has a maximum the distribution density of conditional probability of occurrence $\vec{X}^*$ in *m*-th scenario:

$$\exists! m^* \in C_m(\rho(\vec{X}_m^* \mid R_m))(m = 0..M-1) : \rho(\vec{X}_m^* \mid R_m) \to \max, \quad (2)$$

where $C_m(\rho(\vec{X}_m^* \mid R_m))$ – is the set of m-th indexes of distribution densities the conditional probability of occurrence $\vec{X}^*$ have a precedent in the *m*-th scenario.

In order to correctly solve the problem of classification was conducted factor analysis of the variables of the condition patients.

The factor analysis. Factor analysis based on principal component (PC) analysis was used as a method for aggregating state variables into a complexes, which pairwise would be poorly are correlated variables (the orthogonalization) for the purpose of eventual reducing the dimension of the state variables.

As is well known the variance multivariate random variables is characterized by the correlation matrix Σ.

The essence of the PC method is to decompose a matrix of values of the state variables X° (1):

$$X^{\circ} = TP^{T} = \sum_{\alpha=1}^{A} t_{\alpha} p_{\alpha}^{T} \,, \qquad (3)$$

where T – it's matrix scores with dimension $I \times A$; P – it's matrix loadings with dimension $J \times A$; A – is the number of principal components.

Matrix loadings P is the transition matrix out of the original space of the state variables $X_1..X_J$ (J -dimension) into principal components space (A-dimension).

In the PCA method are used new aggregated variables $t_{\alpha}\ (\alpha = 1..A)$, which are a linear combination of the original variables $X_j\ (j = 1..J)$ [12]: $T = X^{\circ}\left(P^{T}\right)^{-1}$.

Suppose Λ is a matrix of the eigenvalues of the correlation matrix Σ, $\Lambda = diag\{\lambda_1,...\lambda_A\}$.

The matrix unknown variables (PC) F of dimension $I \times A$, where $A < J$, is defined as a linear combination of:

$$F = X^{\circ} L \,, \qquad (4)$$

where L – is a matrix consisting of the projections eigenvectors of matrix Σ, corresponding to the eigenvalues λ_{α} of the matrix Λ, on the axis of the original coordinate system of the state variables space $X_1..X_J$.

In this case, the matrix itself L when $A = J$ by construction is orthogonal: $L^{T} L = LL^{T} = E$, $L^{T} = L^{-1}$, where E – the identity matrix.

The matrix of factor loadings is given by:

$$P = L\Lambda^{\frac{1}{2}} \,. \qquad (5)$$

The rows matrix P correspond to the original variables, the columns correspond to factors (or PC). For the matrix loadings the following relation holds:

$$P^T P = \Lambda = diag\{\lambda_1, \ldots \lambda_A\}, \tag{6}$$

where values $\lambda_1 \geq \ldots \geq \lambda_A \geq 0$ are eigenvalues of the covariance matrix Σ. Expression (3) taking into account (5) can be rewritten as:

$$X^{\circ} = TP^T = T\left(L\Lambda^{\frac{1}{2}}\right)^T. \tag{7}$$

The elements of the matrix scores T they represent normalized PC, which obtained from ordinary PC by multiplying $\Lambda^{-\frac{1}{2}}$: $T = F\Lambda^{-\frac{1}{2}}$, $T = X^{\circ}L\Lambda^{-\frac{1}{2}}$, thus in view of equations (4-7) we have:

$$F = T\Lambda^{\frac{1}{2}} = X^{\circ}P\Lambda^{-\frac{1}{2}}. \tag{8}$$

State recognition. The problem of status recognition is reduced to the classification problem (2). In order to avoid multicollinearity we apply orthogonalization procedure state variables (8). Furthermore to solve the problem of classifying an object's state was used a probabilistic neural network with the following structure (Figure 3):

- the input layer $F_1^* \ldots F_A^*$ – The input elements are the values of the vector projections of observed precedent symptoms $\vec{F}^*$;
- the layer samples $\rho_{1m} \ldots \rho_{Mm}$ – Centers classes of training sample. The number of patterns equal to the count of classes in the training set.

The input layer and the layer samples are fully meshed structure. Activity the element layer samples was determined by the dependence which corresponded to the probability distribution density according to the Student's t- law (which is appropriate for the limited samples): $\rho_{lm} = \rho\left(\vec{F}_m^* \mid R_l\right) = \dfrac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi n}\,\Gamma\left(\frac{n}{2}\right)}\left(1 + \dfrac{t_{ml}^2}{n}\right)^{-\frac{n+1}{2}}$, where Γ – is a the Euler gamma function with *n* degrees of freedom ($n = K_l + K_m - 2$; K_l, K_m – the number of precedents in the corresponding classes, $l, m = 1..M$);

t_{lm} – statistics Student:

$$t_{lm} = \sqrt{\frac{MD_{lm}^2}{\frac{1}{K_l} + \frac{1}{K_m}}}, \qquad (9)$$

where MD_{lm}^2 – Mahalanobis distance from an unknown precedent (is assuming that it belongs to class *l*) to the *m*-th sample – $MD_{lm}^2 = \frac{1}{A}\left(\vec{F}^* - \left\langle \vec{F}_m \right\rangle\right)^T \Sigma_{pooled}^{-1}\left(\vec{F}^* - \left\langle \vec{F}_m \right\rangle\right)$, where $\vec{F}^*$ – the values are projections PC vector of observed symptoms of unknown precedent;

$\left\langle \vec{F}_m \right\rangle$ – average projections values PC vector of the observed symptoms of layer samples element;

Σ_{pooled} – the combined correlation matrix for the considered scenarios (for classes) – is determined in view of belonging of the precedent to one, or a another class: $\Sigma_{pooled} = \frac{1}{K_l + K_m - 2}\left((K_l - 1)\Sigma_l + (K_m - 1)\Sigma_m\right)$;

- The output layer $m^*, \rho(\vec{F}_{m^*}^* \mid R_{m^*})$ (an output element) this is a discriminator threshold values which indicates the element of the fiber samples with a maximum value of activity (i.e. indicates the class to which belongs unknown precedent).

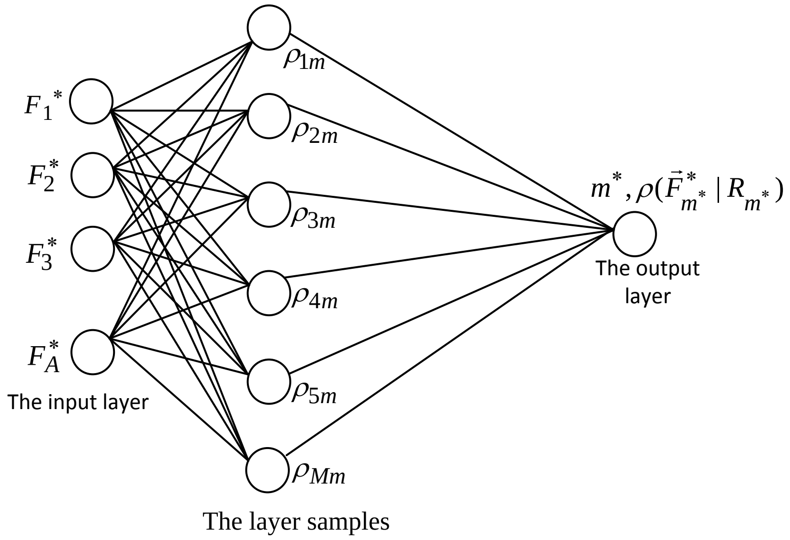

Fig. 3. Architecture of probabilistic neural network.

Should be noted that the values of Student's statistics (9) depend on the choice of the basis, wherein is assessed degree of proximity between the precedent and samples (in solving process the classification problem), as well as between samples (in the analysis of the importance of the distances between classes). Therefore there is a need to structure the additional statistical decision rule choice of a single (supporting) basis. According to the principle of maximum likelihood as the decision rule of choice of the support basis m^* was adopted the following:

$$\exists! m^* \in C_m(t_{lm})(l, m = 0..M-1) : \min_l t_{lm} \to \max, \qquad (10)$$

where $C_m(t_{lm})$ – set of *m*-th indexes statistics.

According to the principle of maximum likelihood as a criterion for the transition from one state to the other state of patients (keeping in mind that they are ordered by the level of disease progression), it can be used a Bayesian decisive rule: $\forall m = 0..M-1 : \dfrac{\rho\left(\vec{F}^*_{m+1} \mid R_{m+1}\right)}{\rho\left(\vec{F}^*_m \mid R_m\right)} \geq 1$

, which is true under the condition that $\frac{P(R_m)}{P(R_{m+1})} \cdot \frac{\rho\left(\vec{F}_{m+1}^*\right)}{\rho\left(\vec{F}_m^*\right)} \approx 1$, where $P(R_m)$ – is a priors probabilities of implementation the classes.

4. Information technology solving classification state problem elements of the biomedical system

System model of information and analytic provision of decision making process by decision-maker-person (DM) in the medico-biological diagnosis of state the patients is shown in Fig. 4. Composition and structure of computer interactive decision support system (CI DSS) for biomedical diagnosis of patients have been formed. The information technology of the calculations in environment which is developed by CI DSS had been developed.

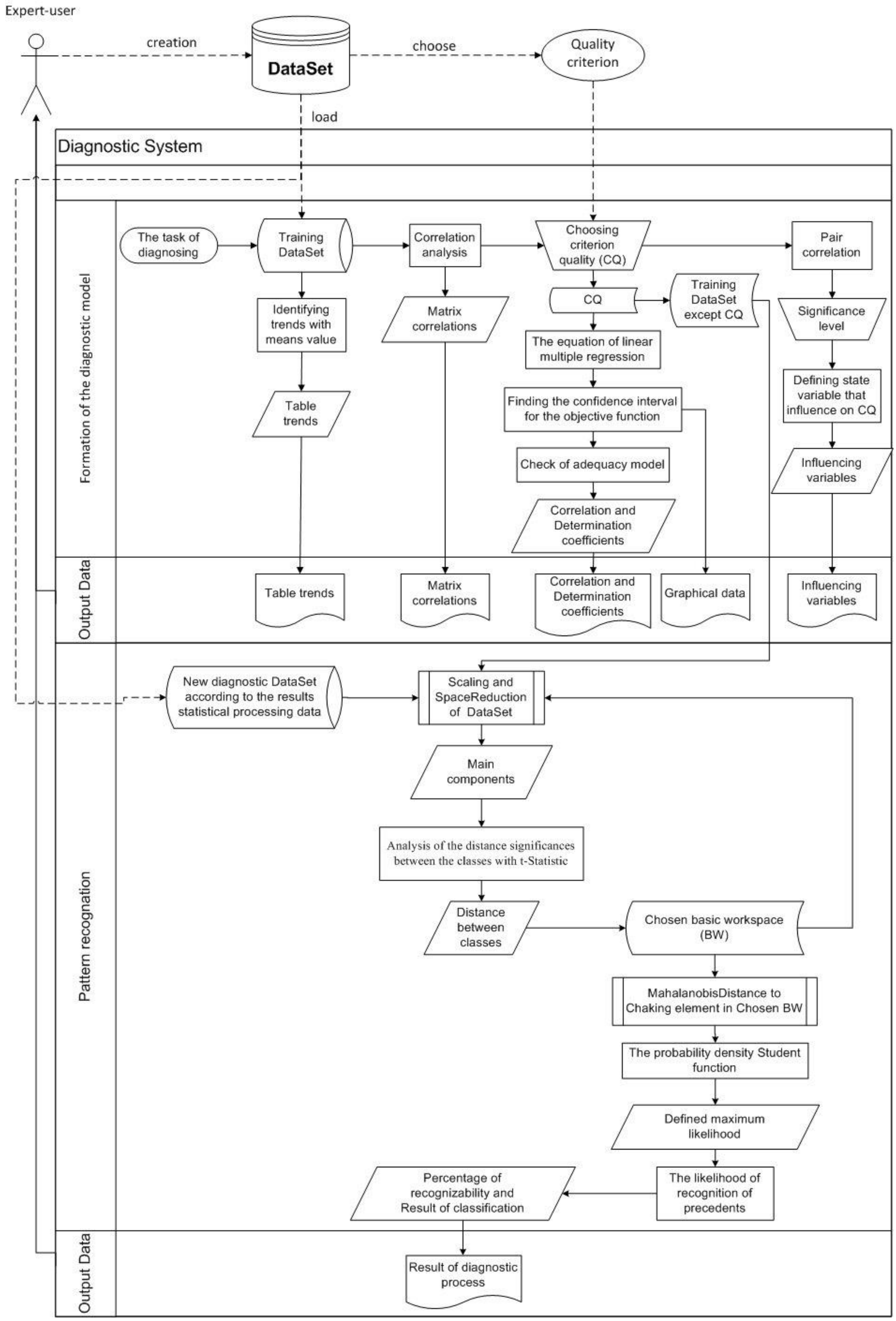

Fig. 4. The system model of information and analytic provision of decision making process by decision-maker-person (DM) in the medico-biological diagnosis of state the patients.

5. The example of the solution classification problem of state elements (patients) of biomedical systems

At the stage of the factor analysis was conducted reduction of the dimension space the PC based on the Kaiser's criterion [12]. It was decided to keep 6 PC of 10 PC, that allowed to reduce the dimension of the state variables.

An analysis of the importance distance between the classes was conducted taking into account selected number of state variables and the number of precedents in the groups (classes). The distance between the centers classes in each of the selected basis spaces is defined based on Student's $\{t_{lm}\}$ statistic. The statistical decision Student's rule had been used – it's a hypothesis of equality of the means [2, 11]. Analysis of the data in Table. IV shows, that the distances between the classes are significant, as evidenced by the excess of the critical values of the statistics Student.

Following the decision rule (10) as a support basis, was selected basis corresponding to the class '0'. Fig. 5 shows the arrangement of classes '0', '1', '2' and '3' in the selected basis space.

Table IV. The results analysis the importance of distances between classes

	Class '0'	**Class '1'**	**Class '2'**	**Class '3'**
These the values of the Student's statistics $\{t_{lm}\}$, which correspond the distance between the centers of the classes when the number PC=6 {critical values – $t_{crt}(n_{min} = 150; P=0,99)=2,61$}				
0	0	23,84	24,04	55,04
1	31,78	0	10,01	21,70
2	34,39	11,74	0	13,14
3	48,90	22,34	9,63	0

The results of solving the problem of classification based on the estimates of probability of belonging precedent to this or that class are shown in the Table. V. The probability of recognizing of the class itself is more than 70%.

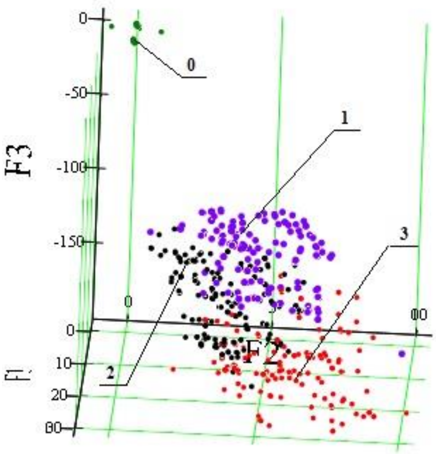

Fig. 5. Location of training set precedents in the space of state variables relative to the support basis.

Table V. The results of solving the classification problem

The probability of recognizing class, %							
Class '0'		**Class '1'**		**Class '2'**		**Class '3'**	
PC=6	*100*	PC =6	*90,3*	PC =6	*71,6*	PC =6	*89,3*

6. The results and conclusions

A system model of the process of diagnosing medical and biological elements of the system (of patients) has been proposed.

Statistical method for classification of state of the elements of dynamic systems based on probabilistic neural network approach has been improved. It now includes the orthogonalization procedure and the reduction of the dimension of the factor space of the state variables and, in contrast to existing methods, uses as a measure of proximity the precedents in the space of Student's t statistic state variables, in the chosen pivotal basis, according to the principle of maximum likelihood. The high result reliability of states precedents recognition achieved if the distance between the classes are significant.

Thus, for example, the results solving problem constructing diagnostic models and state classification of elements biomedical system (of patients) for the particular type of disease us have been obtained. Based on the information content analysis and aggregation variables made the reduction dimension of the space of the state variables.

CI DSS has been developed at the biomedical diagnosis of patients state. There is also offered information technologies calculations in an environment developed by CI DS.

In the current study, mathematical processing variables of the state patients determined their specificity for different stages of lung cancer, that allows you to verify a diagnosis with a probability of at least 70% and to choose an adequate treatment for cancer process when evaluating the aggregate allocated in importance indicators (T stage, Lung lobe, Bombesin, REA, Metastases) by means of developed method and CI DSS that implements it.

Thus, the method of diagnosis of diseases of the respiratory system, which has been developed by us, provides reliable criterias estimate, allows to take into account the mutual influence of different variables on each other and allows you to make the diagnosis of respiratory diseases for each individual patient.

References

1. Rudenko, O.G. Artificial neural network. Textbook / O.G.Rudenko, E.V. Bodyanskiy. – Kharkov: "Company SMITH", 2005. – 408p. (in Russian).
2. Haykin Neural networks: a complete course, second edition/ Haykin, Simon: Trans. from English. Moscow publishing house "Williams", 2006. – 1104 p. (in Russian).
3. Ugryumova, E.M. Trained artificial neural network in the construction of formal mathematical models of systems under a priori uncertainty of data / E.M. Ugryumova // Herald Kharkov University: Coll. Science. Avenue Avg. Mathematical Modelling. Information Technology. Automated control systems. – 2010. – Issue 13 (№890). – P. 237 – 253. (in Russian).
4. Taguchi G. The Mahalanobis–Taguchi Strategy. A pattern technology system / Taguchi G., Jugulum R. New York: John Wiley & Sons, 2002. – 235 p.

5. Vysochina, O.S. Telecommunication network monitoring model based on modified probabilistic neural network / O.S. Vysocina, S.I. Shmatko, A.M. Salman // East European Journal of advanced technologies. – Kharkov, 2010. – №5. – P.67-71. (in Russian).
6. Koltunov, I.A. New statistical methods of pattern recognition and classification for automatic decryption of remote sensing / I.A. Koltunov, A.N. Dumin, V.A. Katrych, R.R. Naumov // Herald Kharkov National University – 2011. – № 966. A series “Radiophysics and Electronics”, Issue 18. – P. 37-49. (in Russian).
7. Kritskyi, O.L. Fisher information matrix for the multivariate method of dynamic conditional correlations DCC-MGARCH(1,1) / O.L. Kritskyi // Herald of Tomsk State University. – 2009. – № 4. A series “Management, computer engineering and Informatics”, Issue 9. – P. 67-82. (in Russian).
8. Igor Loboda, Sergiy Yepifanov, Yakov Feldshteyn, 2012, “A more realistic scheme of deviation error representation for gas turbine diagnostics”, Proceedings of ASME Turbo Expo 2012 GT2012, June 11-15, 2012, Copenhagen, Denmark, 10p.
9. Sokolov O., Radyvonenko O., Korchak T, Gololobova O. Epidemics prediction with the use of neuro-fuzzy methods for time series under uncertainty conditions // Journal of Health Sciences, Poland, 2012. P. 64-70. (in Russian).
10. Afanasyevskaya, V.E. The methodology of solving the problem of classification of technical condition of aircraft gas turbine engine / V.E. Afanasyevskaya, E.M. Ugryumova, V.V. Nerubassky, T.V. Gaidenko // East European Journal of advanced technologies. – 2012. – № 3/10 (57). – P.40-43. (in Russian).
11. Aivazyan, S.A. Applied Statistics and Econometrics basis – Textbook for high schools / S.A. Aivazyan, V.S. Mkhitaryan. – Moscow: UNITY, 1998. – 1022p. (in Russian).
12. Khalafyan, A.A. Textbook STATISTICA 6. Statistical analysis of data / A.A. Khalafyan. – Moscow: Bean-Press Ltd., 2007. – 512 p. (in Russian).