MSC 68T30 DOI: 10.14529/jcem210101

MATHEMATICAL MODELS FOR ANALYZING AND INTERPRETING MICROWAVE RADIOMETRY DATA IN MEDICAL DIAGNOSIS

V. V. Levshinskii, Volgograd State University, Volgograd, Russian Federation, v.levshinskii@volsu.ru.

The present paper is devoted to the improvement and application of mathematical models for the dynamic description of the patient state in the diagnosis of breast cancer and venous diseases based on microwave radiometry data. I present and describe in detail a modified approach for constructing interpretable features in thermometric data. A model evalutation is performed by constructing classification algorithms in the following feature spaces: temperature values, thermometric features, 2nd, 3rd and 4th degree polynomial features. Best algorithms have sensitivity value of 0.892 and specificity value of 0.813 in the mammary glands dataset and sensitivity value of 0.961 and specificity value of 0.925 in the lower extremities dataset. The algorithms built also provide an explanation of result in terms which are understandable for clinicians. The most important features in thermometric data are presented, as well as an example of explanation building.

Keywords: microwave radiometry, feature construction, mathematical modeling

Introduction

Nowadays, the development of intelligent systems based on artificial intelligence methods is a very urgent task. Such technologies can significantly improve the quality of life in a wide variety of fields. For example, in medicine, intelligent systems are able to identify important and subtle details in examination data and thereby improve diagnostic efficiency, as well as reduce experience requirements for specialists performing diagnostics. At the same time, the most interesting are intelligent advisory systems that not only use machine learning methods and algorithms, but also contain mechanisms for explaining the proposed solutions. The development of such systems requires the application of mathematical modeling, data analysis, and machine learning methods.

A promising diagnostic method is the microwave radiometry, which is based on measuring the intrinsic electromagnetic radiation of human tissues in the microwave and infrared wavelength ranges. The method is absolutely safe, allows non-invasive detection of temperature anomalies at a depth of several centimeters and is applied in various fields of medicine [1], including the early diagnosis of breast cancer [2, 3], as well as the diagnosis of venous diseases [4].

The examination technique consists of consecutive measurements of internal (microwave) and surface (infrared, skin) temperatures which are recorded as numerical data and the subsequent search for temperature anomalies in the examination data. The task of finding anomalies in thermometric data is a complex intellectual task requiring long training and years of experience. The interpretation and formalization of expert knowledge, as well as knowledge extraction from data, are the key stages in the development of models for solving such problems.

The research aim is to improve and apply mathematical models for the dynamic description of the patient state in the diagnosis of breast cancer and venous diseases based on microwave radiometry data.

1. Microwave Radiometry in Diagnostics

The microwave radiometry is a biophysical non-invasive examination method, which is based on the consecutive measurements of internal (microwave) and surface (infrared, skin) temperatures at specific points and the subsequent recording of temperatures as numerical data.

A diagnostician is performing an analysis of the data obtained, which can be displayed in the form of thermograms or maps of temperature fields in order to detect temperature anomalies, and makes a conclusion about the patient health state, or the need for further examination by more expensive or dangerous methods. The idea of diagnostic method is that the presence of temperature anomalies indicates the presence of structural changes.

1.1. Breast Cancer

In the early diagnosis of breast cancer, the method allows to effectively detect fast-growing tumors and significantly increase the efficiency of the examination in conjunction with other methods. For example, the combined diagnostic sensitivity together with mammography is 98% [3].

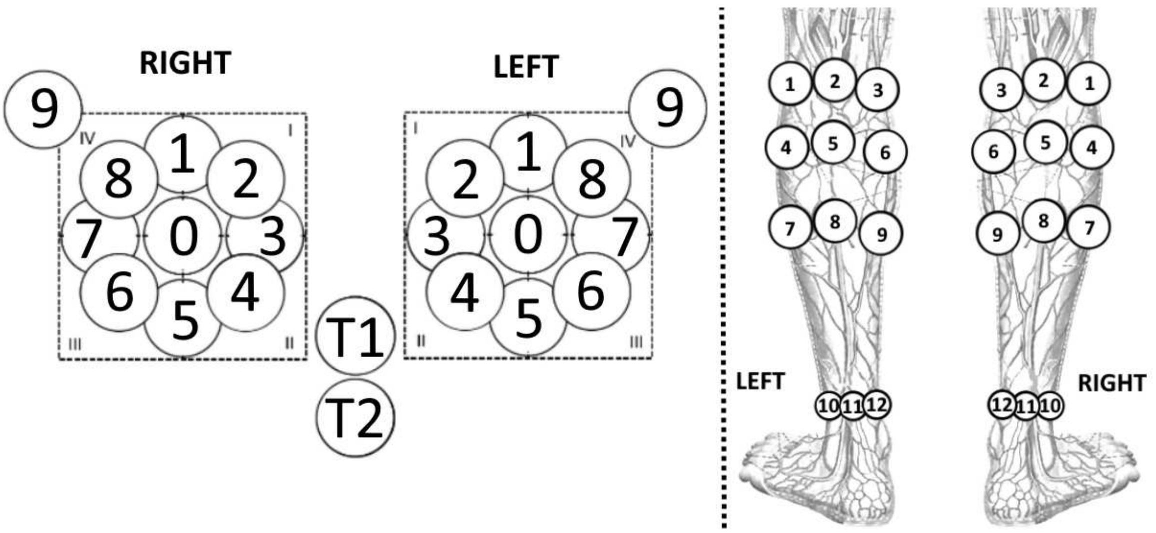

Fig. 1. Sampling points on breasts and legs

The microwave radiometry examination of the mammary glands consists of consecutive measurements of internal and surface temperatures at points $0, \ldots, 8$, axillary region (point 9) and reference points $T1$ and $T2$ according to Figure 1. Methodology assumes that the patient is in supine position, however, in practice, measurements can be additionally carried out in a sitting position.

1.2. Venous Diseases

Microwave radiometry is also applied in the early diagnosis and dynamic monitoring of venous diseases of the lower extremities [4].

The microwave radiometry examination of the lower extremities consists of consecutive measurements of internal and surface temperatures at 12 symmetrical points located on the posterior surface of lower legs according to Figure 1. Two series of measurements are performed for the patient being in different positions: lying on the stomach and standing.

2. Data and Methods

2.1. Datasets

The following two datasets are being considered:

1. thermometric data of 518 mammary glands, among which 166 are healthy and 352 are having various diseases including breast cancer (166);
2. thermometric data of 292 lower extremities, among which 36 are healthy and 256 are having various venous diseases.

Formally, a dataset can be represented as a matrix

$$X = \begin{bmatrix} t_1^1 & t_2^1 & \dots & t_n^1 \\ t_1^2 & t_2^2 & \dots & t_n^2 \\ \dots & \dots & \dots & \dots \\ t_1^m & t_2^m & \dots & t_n^m \end{bmatrix}, \ y = \begin{bmatrix} y_1 \\ y_2 \\ \cdot\,\cdot \\ y_m \end{bmatrix}, \ Y = \{1, 2, \dots, C\}, \tag{1}$$

where m is a number of objects in a dataset, n is a number of features, $x^i = (x_1^i, \dots, x_n^i)$ – the feature vector of the object i, Y is the set of class labels, and $y_i \in Y$ is a label.

Feature vector of the mammary gland contains the values of internal and surface temperatures of points measured according to Figure 1. There are 24 values in total. Let's group the temperatures and denote

$$\begin{aligned} mg^i = (&t_0^{i,mw}, t_1^{i,mw}, \dots, t_9^{i,mw}, t_0^{i,ir}, t_1^{i,ir}, \dots, t_9^{i,ir}, \\ &t_{1,a}^{i,mw}, t_{2,a}^{i,mw}, t_{1,a}^{i,ir}, t_{2,a}^{i,ir}), \end{aligned} \tag{2}$$

where

1. $T^{i,mw} = (t_0^{i,mw}, t_1^{i,mw}, \dots, t_9^{i,mw})$ is a group of internal temperatures of the mammary gland;
2. $T^{i,ir} = (t_0^{i,ir}, t_1^{i,ir}, \dots, t_9^{i,ir})$ is a group of surface temperatures;
3. $T_a^{i,mw} = (t_{1,a}^{i,mw}, t_{2,a}^{i,mw})$ and $T_a^{i,ir} = (t_{1,a}^{i,ir}, t_{2,a}^{i,ir})$ are internal and surface temperatures at the additional points $T1$ and $T2$ respectively.

The subscript indicates the point number. The superscript mw (internal) or st (surface) indicates which type the temperature values belong to.

Feature vector of the lower extremity contains measurements according to Figure 1 in two positions: standing and lying down. There are 48 values in total. Almost similarly, let's group and denote

$$\begin{aligned} lg^i = (&t_1^{i,mw,st}, t_2^{i,mw,st}, \dots, t_{12}^{i,mw,st}, t_1^{i,ir,st}, t_2^{i,ir,st}, \dots, t_{12}^{i,ir,st}, \\ &t_1^{i,mw,ly}, t_2^{i,mw,ly}, \dots, t_{12}^{i,mw,ly}, t_1^{i,ir,ly}, t_2^{i,ir,ly}, \dots, t_{12}^{i,ir,ly}). \end{aligned} \tag{3}$$

The indices st (standing) and ly (lying) have been added here to indicate in which position the patient was when the measurements were taken.

2.2. Feature Construction and Interpretation

An important stage is the actualization and formalization of existing knowledge about the behavior features of the temperature fields. During the search for anomalies, specialists analyze not just temperature values, but their various ratios, "characteristic" features.

Various qualitative features of breast cancer have been identified and described through research and analysis of microwave radiometry data [5]: increased value of thermoasymmetry between the similarly-named points of the mammary glands; increased temperature spreading between individual points in the affected mammary gland; nipple temperature difference; the ratio of surface and internal temperatures and some others.

In the diagnosis of venous diseases lateral-medial and axial gradients [6], which are justified by physiological features, are also important characteristics.

The listed features are a set of qualitative characteristics of temperature anomalies. Further, mathematical descriptions are being offered for each feature [5]. For example, the increased value of thermoasymmetry between the similarly-named points can be described by functions of the form $f = t_{r,i}^{mw} - t_{l,i}^{mw}$, where $t_{r,i}^{mw}$ and $t_{l,i}^{mw}$ are internal temperatures of the i-th point of the right and left glands. There are about 900 such functions, which makes the feature space and the output of algorithms rather cumbersome.

Currently, many of the descriptions are presented in a more general form to minimize the feature space. There are also groups of uniform patterns in the mammary glands and the lower extremities data, while in both cases there is a certain general principle of constructing features, i.e. a general set of universal features in thermometric data. This set of features is represented in the form of hypotheses about the behavior of the temperature fields and the corresponding generalized mathematical descriptions:

1. The hypothesis of an insignificant temperature difference, according to which healthy organs or body parts are characterized, by low values of the following functionals:
   (a) Temperature deviation
   $$F_1(T) = ST_{dev}(T) = \sqrt{\frac{\sum\limits_{t \in T}(t - \overline{T})^2}{|T| - 1}}, \tag{4}$$
   where T is temperatures, $\overline{T}$ is the average value of temperatures in T, $|T|$ is a number of temperature values in T. More specific:
      i. Internal temperatures deviation
      $$f_{mg,1}(mg^i) = F_1(T^{i,mw}). \tag{5}$$
      ii. Similarly for the lower extremities with division into standing/lying measurement position
      $$f_{lg,1}(lg^i) = F_1(T^{i,mw,ly}). \tag{6}$$
   (b) Internal gradients deviation, which are the differences between internal and surface temperatures at the corresponding points. Internal gradients form separate groups, which are defined as element-wise differences
   $$T^{i,g} = T^{i,mw} - T^{i,ir}. \tag{7}$$

The maximum and minimum values, (4), as well as L^p norms are used as characteristics describing the deviation of internal gradients:

$$\begin{aligned} F_2(T) &= \|T\|_1 \,, \\ F_3(T) &= \|T\|_2 \,, \\ F_4(T) &= \|T\|_\infty \,, \end{aligned} \tag{8}$$

where

$$\begin{aligned} \|T\|_p &= (\sum_{t \in T} |t|^p)^{\frac{1}{p}}, \\ \|T\|_\infty &= \max_{t \in T} |t|. \end{aligned} \tag{9}$$

More specific, the maximum absolute value of internal gradients of the mammary gland

$$f_{mg,2}(mg^i) = F_4(T^{i,g}). \tag{10}$$

(c) Temperature deviation from average, temperature oscillation and others.

2. The hypothesis about the symmetry of the temperature fields of paired organs (body parts), according to which healthy paired organs are characterized by a slight deviation of temperatures at the corresponding points (subregions), as well as a slight difference in related characteristics. The following characteristics are used as generalized measures of symmetry

$$\begin{aligned} F(T_c, T_p) &= \|T_c - T_p\| \,, \\ F(T_c, T_p) &= \|T_c\| - \|T_p\| \,, \end{aligned} \tag{11}$$

where $\|z\|$ is a functional, $T_c - T_p$ is element-wise difference, T_c is "current" and T_p is "paired" group of temperatures. These characteristics require an additional step of data preprocessing, as well as the existence of a pair for each object in the sample. For example, in the process of preprocessing lower extremities dataset, if the left extremity is being considered, then the "current" temperature group is internal or surface temperatures of the left extremity, and its "paired" group will be the internal or surface temperatures of the right extremity, respectively.

For paired temperature groups characteristics are mainly based within the framework of the previous hypothesis:

(a) The maximum absolute value of temperature difference between the similarly-named points

$$F_5(T_c, T_p) = F_4(T_c - T_p). \tag{12}$$

(b) Difference between standard deviations of temperatures

$$F_6(T_c, T_p) = F_1(T_c) - F_1(T_p). \tag{13}$$

(c) Difference between average values and others

$$F_7(T_c, T_p) = \overline{T_c} - \overline{T_p}. \tag{14}$$

3. The temperature stability hypothesis, according to which healthy organs or body parts are characterized by insignificant differences in temperatures measured at different positions of the patient. Features of this group characterize the degree of proximity of temperature fields in different positions and are practically similar to features determined within the framework of the symmetry hypothesis.

   For instance, the following features:

   (a) Difference of average surface temperatures measured in standing and lying positions

   $$f_{lg,2}(lg^i) = F_7(T^{i,ir,st}, T^{i,ir,ly}) = \overline{T^{i,ir,st}} - \overline{T^{i,ir,ly}}. \tag{15}$$

   (b) Maximum absolute value of the difference between the internal temperature gradients measured in the standing and lying positions

   $$f_{lg,3}(lg^i) = F_5(T^{i,g,st}, T^{i,g,ly}) = \left\| T^{i,g,st} - T^{i,g,ly} \right\|_\infty. \tag{16}$$

   These features are constructed in the lower extremitites data. For the mammary glands, there are no measurement data in several positions in the sample.

4. Hypotheses related to the physiological structure of organs (body parts). For instance, the difference between nipple temperatures in breasts data

   $$F_8(T_c, T_p) = T_{0,c} - T_{0,p}, \tag{17}$$

   deviation of temperature values relative to the point 9, gradients of additional points in breasts data or the values of lateral-medial and axial gradients in the lower extremities data and others.

In this way the feature space can be redefined. For each object in the dataset the function values f are calculated. Sixty-five new features are constructed in the mammary glands dataset, and 128 in the lower extremitites dataset. Further, by binarizing [7] the obtained values, the construction of the set of thermometric features is performed

$$S = (\phi_1, \phi_2, \ldots, \phi_s), \tag{18}$$

where s is a number of features.

Thermometric feature is a triple $\phi = (f, I, W)$, where I is an interval and W is a weight (informativity f by I), or a quantitative indicator that determines how well the characteristic separates objects of one class from other classes. Thermometric feature is observed in the object x^i, if $f(x^i) \in I$. A key feature of thermometric characteristic is interpretability. This fact follows from hypotheses about the behavior of temperature fields, which, in turn, evolve from qualitative features.

Vector of values of thermometric features $(\phi_1(x^i), \phi_2(x^i), \ldots, \phi_s(x^i))$ will describe the state of object i. The element of the vector with the index j will be equal to 1 if the feature j is observed in the object x^i and 0 otherwise.

Thermometric features are the basic building blocks for more complex structures, for instance, 2-dimensional features [8], and classification models.

2.3. Classification

Based on thermometric features, it is possible to construct various classification algorithms. To evaluate their effectiveness, a weighted voting algorithm is constructed, a key feature of which is the possibility of explaining the diagnostic result.

Consider a binary classification problem. Label 0 corresponds to a class "Healthy" and label 1 corresponds to class "Sick". The classification algorithm is defined as

$$a(x^i) = \begin{cases} 1, \text{if } h_W(x^i) \geq 0.5, \\ 0, \text{otherwise}, \end{cases} \tag{19}$$

where

$$h_W(x^i) = g(W_0 + \sum_{j=1}^{s} W_j \phi_j(x^i)) \tag{20}$$

is the sum of weights of thermometric features, W_j is a weight of feature ϕ_j, and $g(z)$ is a sigmoid.

The construction of classification model consists of the following steps:

1. Distinguish temperature groups, perform feature construction and binarization, find thermometric features (18);
2. Transform the data into a binary matrix X' whose element at the intersection of the i-th row and j-th column is 1 if thermometric feature j is observed in object i, and 0 otherwise;
3. Weigh up and select the most effective thermometric features by logistic regression with L_1-regularization, in which case the weights of insignificant features are zeroed [9]. A classification model is constructed for X'.

2.4. Modeling Exercise

To evaluate the effectiveness of thermometric features, several classification algorithms have been built using the logistic regression method. For comparison, algorithms were built in the following feature spaces: temperature values, values of thermometric functions, thermometric features, 2nd, 3rd and 4th degree polynomial features [10].

The efficiency of the algorithm was evaluated by nested cross-validation with preservation of class balance. The number of blocks on the external level is 9, on the internal level is 8. The advantage of nested cross-validation is that the evaluation of the algorithm, which requires pre-tuning of parameters (e.g., regularization coefficient), is always performed on the data unknown during training, and therefore is fair enough [10].

The G-measure [11] was used as an evaluation metric, which is determined by the formula

$$G_{mean} = \sqrt{Sens \cdot Spec}, \tag{21}$$

where

$$Sens = \frac{TP}{TP + FN} \tag{22}$$

is a sensitivity and

$$Spec = \frac{TN}{TN + FP} \tag{23}$$

is a specificity, TP is a number of true positives, FN - false negatives, TN - true negatives, FP - false positives. Sensitivity and specificity are traditional measures of the effectiveness of diagnostic methods, and the G-measure is a fairly fair estimate for unbalanced samples.

3. Results and Discussion

Classification results for the mammary glands dataset are presented in Table 1.

Table 1

Mammary glands classification performance

Feature space	G_{mean}		$Sens$		$Spec$	
	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
Thermometric features	0.85	0.043	0.892	0.051	0.813	0.073
Thermometric features (values)	0.811	0.047	0.801	0.061	0.824	0.08
Temperature values	0.778	0.039	0.747	0.035	0.813	0.067
2nd degree polynomial features	0.78	0.045	0.75	0.048	0.812	0.072
3rd degree polynomial features	0.793	0.044	0.77	0.043	0.818	0.061
4th degree polynomial features	0.804	0.047	0.798	0.046	0.812	0.087

The highest sensitivity and overall classification performance is achieved using thermometric features. The deviation of sensitivity for all algorithms is about 0.05, specificity - 0.07. Increasing the order of the polynomial features increases the sensitivity and overall performance of the algorithm. The highest specificity is achieved using values of thermometric functions. The algorithm that classifies only by temperature values has the lowest sensitivity.

Classification results for the lower extremities dataset are presented in Table 2.

Table 2

Lower extremities classification performance

Feature space	G_{mean}		$Sens$		$Spec$	
	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
Thermometric features	0.939	0.078	0.961	0.046	0.925	0.139
Thermometric features (values)	0.838	0.07	0.816	0.045	0.869	0.143
Temperature values	0.519	0.2	0.578	0.098	0.525	0.233
2nd degree polynomial features	0.577	0.136	0.586	0.096	0.588	0.22
3rd degree polynomial features	0.582	0.243	0.609	0.08	0.625	0.294
4th degree polynomial features	0.614	0.159	0.629	0.089	0.619	0.245

Here, the highest performance of classification is achieved when applying thermometric features. At the same time, the difference between scores is remarkable. For algorithms based on temperature values, there is a significant deviation of performance arising from the deviation of specificity (about 0.25). The sensitivity score is quite stable and is about 0.08. The worst result is achieved when constructing a classifier based on temperature values only. Polynomial features increase the efficiency of the classification.

Tables 3 and 4 contain top 5 features with the highest absolute values of weights obtained after training the weighted voting classifier on breasts and legs datasets respectively. In both cases thermal asymmetry features have the most weight and other groups of features are less common. However, all groups of features are important for achieving high performance.

Table 3

Top 5 features (breasts dataset) sorted by absolute value of weight

Feature	W	Sens	Spec
$\left\lVert T_c^{i,ir} - T_p^{i,ir} \right\rVert_2 \in (2.166, 2.215)$	-6.123	0.0	0.96
$ST_{dev}(T_c^{i,g} - T_p^{i,g}) \in (0.27, 0.292)$	4.699	0.06	0.99
$\left\lVert T_c^{i,ir} - T_p^{i,ir} \right\rVert_2 \in (2.495, 2.696)$	2.651	0.11	1.0
$\left\lVert T_c^{i,mw} - T_p^{i,mw} \right\rVert_2 \in (1.14, 1.179)$	-2.624	0.01	0.94
$\left\lVert T_c^{i,ir} - T_p^{i,ir} \right\rVert_2 \in (1.98, 2.087)$	2.577	0.03	1.0

Table 4

Top 5 features (legs dataset) sorted by absolute value of weight

Feature	W	Sens	Spec
$ST_{dev}(T_c^{i,g,st} - T_p^{i,g,st}) \in (0.394, 0.414)$	-1.542	0.02	0.78
$ST_{dev}(T_c^{i,mw,ly} - T_p^{i,mw,ly}) \in (0.287, 0.389)$	1.243	0.23	1.0
$\left\lVert T^{i,g,ly} \right\rVert_1 \in [0, 19.05)$	1.183	0.14	1.0
$\left\lVert T_c^{i,g,ly} - T_p^{i,g,ly} \right\rVert_\infty \in [0, 0.55)$	-1.164	0.01	0.56
$\left\lVert T_c^{i,ir,ly} - T_p^{i,ir,ly} \right\rVert_2 \in (1.581, \infty)$	1.025	0.8	0.89

It should be noted that each feature can be interpreted. For instance, in Table 4 features can be interpreted as the following:

1. Normal deviation of the difference between temperature gradients of the lower extremities, measured in standing position;
2. Increased deviation of the difference internal temperatures of the lower extremities, measured in the supine position;
3. Suspicious deviation of internal gradients of the lower extremities, measured in the supine position;
4. Normal maximum absolute value of differences between internal gradients of the lower extremitites, measured in the supine position;
5. Increased deviation of surface temperature differences between the lower extremities, measured in the supine position.

A set of descriptions of the observed features forms an explanation of decision.

The presented mathematical models of the patient state in diagnosis based on microwave radiometry data and the constructed feature space are applied to solve the binary problem of diagnosing breast diseases, as well as diagnosing venous diseases. The constructed feature spaces showed not only their effectiveness, but also the possibility of explaining the result.

The reported study was funded by RFBR, project number 19-31-90153.

References

1. Goryanin I., Karbainov S., Shevelev O., Tarakanov A., Redpath K., Vesnin S., Ivanov Y. Passive Microwave Radiometry in Biomedical Studies. *Drug Discovery Today*, 2020, vol. 25, issue 4, pp. 757–763. DOI: 10.1016/j.drudis.2020.01.016
2. Zamechnik T. V., Losev A. G., Petrenko A. Yu. Guided Classifier in the Diagnosis of Breast Cancer According to Microwave Radiothermometry. *Mathematical Physics and Computer Simulation*, 2019, vol. 22, no. 3, pp. 52–66. (in Russian). DOI: 10.15688/mpcm.jvolsu.2019.3.5.
3. Vesnin S., Turnbull A., Dixon J., Goryanin I. Modern Microwave Thermometry for Breast Cancer. *Journal of Molecular Imaging & Dynamics*, 2017, vol. 7, issue 2, 1000136. DOI: 10.4172/2155-9937.1000136
4. Zamechnik T. V., Larin S. I., Losev A. G. [*Combined Radiothermometry as a Method for Studying Venous Circulation of the Lower Extremities: Monograph*]. Volgograd, VolgGMU Publishing House, 2015. (in Russian)
5. Losev A. G., Levshinskii V. V. The Thermometry Data Mining in the Diagnostics of Mammary Glands. *UBS*, 2017, issue 70, pp. 113–135. (in Russian)
6. Anisimova E. V., Zamechnik T. V., Losev A. G., Mazepa E. A. Some Characteristic Signs in Diagnostics of Venous Diseases of Lower Extremities by the Method of Combined Thermography]. *Journal of New Medical Technologies*, 2011, vol. XVIII, no. 2, pp. 329–330. (in Russian)
7. Vorontsov K. V. [*Lectures on Logical Classification Algorithms*], 2007, available at: http://www.ccas.ru/voron/download/LogicAlgs.pdf (accessed on February 25, 2021). (in Russian)
8. Zenovich A. V., Baturin N. A., Medvedev D. A., Petrenko A. Yu. Algorithms for the Formation of Two-Dimensional Characteristic and Informative Signs of Diagnosis of Diseases of the Mammary Glands by the Methods of Combined Radio Thermometry. *Mathematical Physics and Computer Simulation*, 2018, vol. 21, no. 4, pp. 44–56. (in Russian). DOI: 10.15688/mpcm.jvolsu.2018.4.4
9. Flach P. *Machine Learning: The Art and Science of Algorithms that Make Sense of Data*. Cambridge, Cambridge University Press, 2012.
10. Raschka S., Mirjalili V. *Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow 2, 3rd Edition.* Birmingham, Packt Publishing, 2019.
11. Bekkar M., Djema H., Alitouche T. A. Evaluation Measures for Models Assessment over Imbalanced Data Sets. *Journal of Information Engineering and Applications*, 2013, vol. 3, no. 10, pp. 27–38.

Vladislav V. Levshinskii, Postgraduate Student at the Department of Mathematical Analysis and Function Theory, Volgograd State University (Volgograd, Russian Federation), v.levshinskii@volsu.ru.

Received February 27, 2021.

УДК 004.89 **DOI: 10.14529/jcem210101**

МАТЕМАТИЧЕСКИЕ МОДЕЛИ ДЛЯ АНАЛИЗА И ИНТЕРПРЕТАЦИИ ДАННЫХ МИКРОВОЛНОВОЙ РАДИОТЕРМОМЕТРИИ В МЕДИЦИНСКОЙ ДИАГНОСТИКЕ

В. В. Левшинский

Работа посвящена доработке и применению математических моделей для динамического описания состояния пациентов по данным микроволновой радиотермометрии в диагностике рака молочных желез и венозных заболеваний. Представлен и подробно описан модифицированный подход для конструирования интерпретируемых признаков в термометрических данных. Для оценки модели выполнено построение алгоритмов классификации для разных признаковых пространств: только температурные данные, термометрические признаки, а также полиномиальные признаки различных порядков. В задаче классификации желез достигнута чувствительность 0.892 и специфичность 0.813, а в задаче классификации голеней – 0.961 и 0.925 соответственно, при этом обеспечивается обоснование решения в терминах, понятных врачу-диагносту. Представлены наиболее значимые закономерности в данных, а также пример построения обоснования.

Ключевые слова: микроволновая радиотермометрия; конструирование признаков; математическое моделирование.

Литература

1. Goryanin, I. Passive Microwave Radiometry in Biomedical Studies / I. Goryanin, S. Karbainov, O. Shevelev, A. Tarakanov, K. Redpath, S. Vesnin, Y. Ivanov // Drug Discovery Today. – 2020. – V. 25, iss. 4. – P. 757–763.

2. Замечник, Т. В. Управляемый классификатор в диагностике рака молочной железы по данным микроволновой радиотермометрии / Т. В. Замечник, А. Г. Лосев, А. Ю. Петренко // Математическая физика и компьютерное моделирование. – 2019. – Т. 22, № 3. – С. 53–67.

3. Vesnin, S. Modern Microwave Thermometry for Breast Cancer / S. Vesnin, A. Turnbull, J. Dixon, I. Goryanin // Journal of Molecular Imaging & Dynamics. – 2017. – V. 7, iss. 2. – 1000136.

4. Замечник, Т. В. Комбинированная радиотермометрия как метод исследования венозного кровообращения нижних конечностей / Т. В. Замечник, С. И. Ларин, А. Г. Лосев. – Волгоград: Изд-во ВолгГМУ, 2015.

5. Лосев, А. Г. Интеллектуальный анализ термометрических данных в диагностике молочных желез / А. Г. Лосев, В. В. Левшинский // Управление в медико-биологических и экологических системах. – 2017. – Вып. 70. – С. 113–135.

6. Анисимова, Е. В. О некоторых характерных признаках в диагностике венозных заболеваний нижних конечностей методом комбинированной термографии / Е. В. Анисимова, Т. В. Замечник, А. Г. Лосев, Е. А. Мазепа // Вестник новых медицинских технологий. – 2011. – Т. XVIII, № 2. – С. 329–330.

7. Воронцов, К. В. Лекции по логическим алгоритмам классификации [Электронный ресурс]. – URL: http://www.ccas.ru/voron/download/LogicAlgs.pdf (дата обращения: 25.02.2021).

8. Зенович, А. В. Алгоритмы формирования двумерных признаков диагностики заболеваний молочных желез методами комбинированной радиотермометрии / А. В. Зенович, Н. А. Батурин, Д. А. Медведев, А. Ю. Петренко // Математическая физика и компьютерное моделирование. – 2018. – Т. 21, № 4. – С. 44–56.

9. Flach, P. Machine Learning: The Art and Science of Algorithms that Make Sense of Data. – Cambridge: Cambridge University Press, 2012.

10. Raschka, S. Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow 2, 3rd Edition / S. Raschka, V. Mirjalili. – Birmingham: Packt Publishing, 2019.

11. Bekkar, M. Evaluation Measures for Models Assessment over Imbalanced Data Sets / M. Bekkar, H. Djema, T. A. Alitouche // Journal of Information Engineering and Applications. – 2013. – V. 3, № 10. – P. 27–38.

Левшинский Владислав Викторович, аспирант кафедры математического анализа и теории функций, Волгоградский государственный университет (г. Волгоград, Российская федерация), v.levshinskii@volsu.ru.

Поступила в редакцию 27 февраля 2021 г.