Nonlinear Dynamics and Neuroscience

Izvestiya Vysshikh Uchebnykh Zavedeniy. Applied Nonlinear Dynamics. 2024;32(3)

Article DOI: 10.18500/0869-6632-003098

Using machine learning algorithms to determine the emotional maladjustment of a person by his rhythmogram*

S. V. Stasenko[1]✉, *O. V. Shemagina*[2], *E. V. Eremin*[1], *V. G. Yakhno*[2], *S. B. Parin*[1], *S. A. Polevaya*[1]

[1]National Research Lobachevsky State University of Nizhny Novgorod, Russia
[2]Federal Research Center A. V. Gaponov-Grekhov Institute of Applied Physics of the Russian Academy of Sciences, Nizhny Novgorod, Russia
E-mail: ✉stasenko@neuro.nnov.ru, olgashemagina@gmail.com, eugenevc@gmail.com, yakhno@appl.sci-nnov.ru, parins@mail.ru, s453383@mail.ru
Received 10.10.2023, accepted 8.11.2023, available online 15.03.2024, published 31.05.2024

Abstract. The *purpose* of this study is to explore the feasibility of identifying emotional maladjustment using machine learning algorithms. *Methods.* Electrocardiogram data were gathered using an event-telemetry approach, employing a software and hardware setup comprising a compact wireless ECG sensor (HxM; Zephyr Technology, USA) and a smartphone equipped with specialized software.For constructing the classifier, the following algorithms were employed: logistic regression, easy ensemble, and gradient boosting. The performance of these algorithms was assessed using the f1 metric. *Results.* It is demonstrated that employing dynamic spectra of the original signals enhances the classification accuracy of the model compared to using the original rhythmograms. *Conclusion.* A method is proposed for automatically determining the level of emotional maladaptation based on an individual's cardiorhythmogram. Information from a portable heart sensor, worn by an individual, is transmitted via Bluetooth to a mobile device. Here, the level of emotional maladaptation is assessed through a pre-trained neural network algorithm. When considering a neural network algorithm, it is recommended to employ a classifier trained on spectrograms.

Keywords: machine-learning algorithms, electrocardiogram, emotional disadaptation, data analysis.

Acknowledgements. The work in terms of collecting and data preprocessing was supported by the Russian Ministry of Science and Education project number 075-15-2021-634, the work in term of data analysis was supported by the frames of the Governmental Project of the Institute of Applied Physics RAS, project No. FFUF-2021-0014, the work in terms of developing the conceptual scheme of the experiment was supported by a grant from the Russian Science Foundation (project number 22-18-20075).

For citation: Stasenko SV, Shemagina OV, Eremin EV, Yakhno VG, Parin SB, Polevaya SA. Using machine learning algorithms to determine the emotional maladjustment of a person by his rhythmogram. Izvestiya VUZ. Applied Nonlinear Dynamics. 2024;32(3):394–404. DOI: 10.18500/0869-6632-003098

*The paper presents materials of a talk given at the conference «Nonlinear dynamics in cognitive research — 2023».

Introduction

The COVID-19 pandemic has highlighted the connection between pathological conditions and the increased risk of cognitive impairment, which includes SARS, chronic stress, inflammatory syndromes, and coagulation disorders [1]. Subsequently, it was discovered that the SARS-CoV-2 virus can induce significant systemic health alterations in patients, resulting in a post-COVID syndrome characterized by a broad spectrum of symptoms and varying durations [2,3]. Frequently reported long COVID symptoms include fatigue, memory issues, breathlessness, sleep disturbances, headaches, loss of taste or smell, muscle weakness, fever, cognitive dysfunction, and mental health challenges [4–7]. The development of new methods and approaches to the rapid diagnosis of chronic stress is an urgent task, considering the current epidemiological (Covid-19) situation [8–10]. Psychological stress plays a pivotal role in the development of numerous physical and neurological diseases. The term "stress" is commonly used to refer to both a potent adverse physical and/or psychogenic external environmental impact, and to a state of psychophysiological stress that emerges under such influences, initially aiding a person in adapting to new environmental conditions. Chronic stress, as a long-term psychophysiological burden, can trigger the manifestation or exacerbation of disease symptoms, act as a risk factor, or worsen the severity of the disease. Emotional overstrain reduces an individual's productivity and the quality of work performed. The clinical implications of chronic emotional stress intersect with neuropsychic anxiety and depressive disorders, significantly diminishing the quality of life for individuals [11]. Primarily, chronic emotional stress has adverse effects on health, and indirectly induces unfavorable endocrine, neuromuscular, and autonomic changes [12]. Daily mental stress underlies many prevalent and serious illnesses, including hypertension, strokes, heart attacks, cancer, and more.

While the cause of acute mental stress is primarily linked to unexpected negative external influences and life changes, the development of chronic stress is largely influenced by a person's personal characteristics and the inadequate functioning of their psychological adaptation mechanisms. The initial step to overcoming chronic emotional stress involves an individual acknowledging that they are in a state of mental overstrain. Emotions are subjective, and diagnosing them depends on a person's ability to accurately comprehend and articulate them. However, this ability isn't equally developed in all individuals [13]. Therefore, establishing a system for swiftly assessing emotional disadaptation in everyday life holds significance in promptly diagnosing emotional overstrain and exhaustion. Providing individuals with information about their current biological state enables timely feedback, indicating the level of their mental stress. This feedback allows them to temporarily alleviate this stress by engaging in physical activity or other pursuits. Special attention is required for individuals who have limited awareness of their emotional state's nuances.

1. Methods

The task of diagnosing a person's psychophysiological state and the extent of emotional disadaptation, based on the user's physiological data, is addressed through the method of recording emotional disadaptation using the cardiorhythmogram. This method involves employing a mobile ECG sensor, with the sensor's data transmitted to a mobile application. The application utilizes a neural network algorithm that was previously trained during the calibration phase to automatically classify cardiorhythmogram data based on the level of disadaptation (refer to Figure 1). In this developed system, the markers for emotional disadaptation are versatile and don't necessitate individual calibration for each user. This holds true as long as the algorithm for classifying RR intervals according to disadaptation levels has been pre-trained using a database containing an ample number of distinct user records.

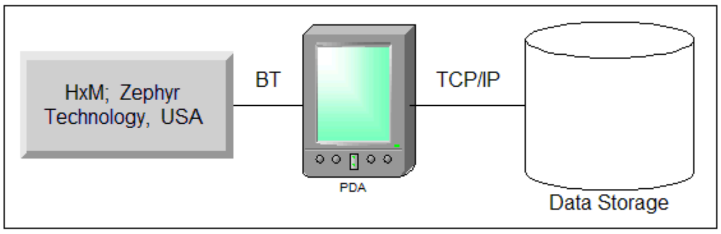

Fig. 1. Diagram of the System for Recording Emotional Disadaptation Levels through Cardiorhythmogram

During the calibration phase, a database is compiled, encompassing sequences of RR intervals derived from ECG data of distinct users. This data is concurrently collected alongside the completion of a questionnaire containing four sets of verbal descriptors delineating emotions based on valence (positive/negative) and activity level (tension/relaxation). These descriptors correspond to four fundamental personal needs:

- Security,
- Independence,
- Achievement,
- Unity.

1.1. Methodology for measuring the level of emotional disadaptation. The dynamics of the psychophysiological system are assessed using a parallel monitoring scheme that involves tracking parameters of autonomic regulation through electrocardiogram data and conducting tests based on a methodology for determining the level of emotional disadaptation (UED) [14,15]. Next, we will describe the methodology in greater detail.

The method for assessing an individual's emotional disadaptation is based on presenting the patient with four groups of verbal characteristics that reflect various emotional states and the degree of their expression. The analysis involves the verbal characteristics chosen by the individual and their assessment using a point scale. This method includes the presentation of four groups of adjectives that reflect the degree of satisfaction of four fundamental personal needs: security, independence, achievement, and unity-closeness. To evaluate the degree of satisfaction of the basic personal need for security, the following system of adjectives and points is employed:

- "calm, peaceful, serene" — 0 points;
- "wary, worried, excited" — 1 point;
- "anxious, frightened, frightened" — 2 points;
- "tortured, tormented, despairing" — 3 points.

To assess the degree of satisfaction of the basic personal need for independence, the following system of adjectives and points is used:

- "lightened, liberated, liberated" — 0 points;
- "hot, indignant, angry" — 1 point;
- "preoccupied, overloaded, overstressed" — 2 points;
- "depressed, oppressed, constrained" — 3 points.

For the basic personal need of achievement, the following system of adjectives and points is applied:

- "satisfied, joyful, proud" — 0 points;
- "motivated, inspired, inspired" — 1 point;
- "overexcited, frantic, inflated" — 2 points;
- "exhausted, devastated, indifferent" — 3 points.

To assess the degree of satisfaction of the basic personal need for unity, the following system of adjectives and points is used:

- "delighted, pleased, prosperous" — 0 points;
- "interested, enthusiastic, admiring" — 1 point;
- "upset, hurt, disappointed" — 2 points;
- "abandoned, lonely, sad" — 3 points.

Each basic personal need is evaluated using a circular scale divided into four equal quadrants intersected by two perpendicular lines at a 45° angle to the horizontal. At each intersection of these lines with the circle, three adjectives are placed, reflecting different degrees of satisfaction of the basic personal need under examination. When the individual selects words that correspond to their condition, a mark is made on the circular scale. The mark may either align with the group of words chosen or, when two different groups of words are selected, fall on a circle or segment between those groups.

Based on the location of this mark, the number of points scored by the individual for each of the four basic personal needs is determined as follows: 1) If the mark is on the border between quadrants, the higher of the two values is assigned; 2) If the mark is in the center of the circle, it is excluded from the analysis.

Then, the average score for all four basic personal needs is calculated. The degree of emotional maladjustment is judged based on this average score:

- 0 points: No emotional disadaptation, physiological relaxation;
- 1 point: Mild emotional disadaptation, physiological stress;
- 2 points: Moderately expressed emotional disadaptation, pathological tension;
- 3 points: Pronounced emotional disadaptation, pathological relaxation.

Telemetric measurements of the electrocardiogram were conducted using a software and hardware complex comprising a compact wireless ECG sensor (HxM; Zephyr Technology, USA) along with a smartphone equipped with specialized software (refer to Figure 1).

Based on the gathered data from different users or an individual user, the algorithm for classifying the sequence of RR intervals was trained, incorporating labels denoting the user's degree of emotional disadaptation.

The pre-trained classification algorithm is subsequently applied to novel RR interval sequence data to autonomously ascertain the level of disadaptation, eliminating the need for a questionnaire.

1.2. Data for constructing a classifier. In constructing the classifier, data obtained from users' heart rhythmograms were utilized (refer to Figure 2). Each data entry corresponded to the level of emotional disadaptation, as determined through a questionnaire. The compiled database encompasses a total of 2222 distinct records.

Each entry consists of a collection of files with the following content:

- rr_filter.csv — Depicts the correlation between the RR-interval value (*RR_filter*) and time (in milliseconds) within the 5-minute period before the test for emotional disadaptation level determination. This file also includes the time of test completion,
- info.csv — Contains data such as the start time of recording (*time*), post ID (*session_id*), subject identifier (*person_id*), age (*old*) and the gender of the subject (*gender*),
- uad.csv — encompasses time stamps for the test's initiation (*ms_begin*) and conclusion (*ms_end*) measured in milliseconds from the recording's start time), the test result indicating varying degrees of satisfaction regarding the four fundamental personal needs—safety, independence, achievement, unity-closeness (U1; U2; U3; U4), the subjective emotional disadaptation level, and the average satisfaction level of the four fundamental personal needs (*medium*).

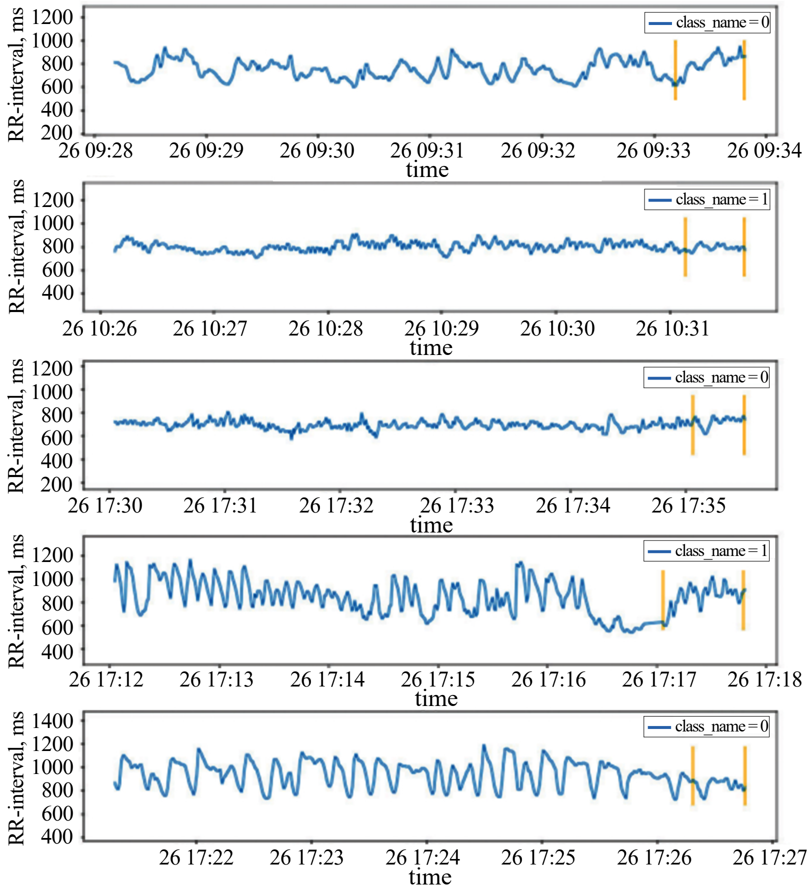

Fig. 2. Examples of cardiograms are provided. The yellow markers signify the interval during which the test is conducted. The graph indicates the classification belonging to a specific class (*class_ name*) (color online)

The time required for subjects to complete the test varied, leading to differing durations of the recorded data. This variability could be inconvenient for the algorithms used in classifier construction. To address this, all records were segmented into sections of 100 samples, commencing from the end of each record. This approach was chosen as the segments coinciding with the time of questionnaire completion were deemed the most informative. Fragments of records containing fewer than 100 samples were disregarded.

All data were standardized and classified into two categories: records featuring an emotional disadaptation level of 0 or 1, as determined by questionnaire results, were categorized as class 0, signifying an absence of disadaptation. Conversely, the remaining records belonged to class 1, indicating the presence of emotional disadaptation.

1.3. Machine-learning algorithms. The subsequent algorithms were employed for constructing the classifiers.

1. Logistic regression [16] — is a method for constructing a classifier based on linear models of the following type: $\hat{y}(w, x) = w_0 + w_1 x_1 + ... + w_p x_p$, where $\hat{y}(w, x)$ is the predicted value. The task of training a linear classifier is to adjust the weight vector w based on the sample X^m. In logistic regression, for this, the problem of empirical risk minimization is solved with a special type of loss function: $\min_{w,c} \frac{1}{2} w^T w + C \sum_{i=1}^{n} \log(\exp(-y_i(X_i^T w + c)) + 1)$. To address the optimization problem, the Broyden–Fletcher–Goldfarb–Shanno algorithm is employed.

2. Easy ensemble [17] — the concept involves training an ensemble of classifiers, each independently determining the classification of an object into a specific class. The ultimate decision is reached through a voting process. This algorithm is particularly applicable to unbalanced data. Training the classifiers involves the entire minority class along with a subset of the majority class. The key parameters of the ensemble include the number of learners (set to 10 in our case) and the algorithm used for training each classifier within the ensemble (we use AdaBoost [19]).
3. Gradient boosting [18] — the concept involves amalgamating multiple weak classifiers that rely on decision trees to form a robust classifier.

These algorithms are implemented in the Scikit-learn library [20] and the Imbalanced-learn library [21].

2. Results

80% of the generated database was utilized for training the classifiers. The remaining 20% constituted the test sample. During the testing phase, the records requiring classification were segmented into segments of 100 samples each. For each segment, the classifier rendered a decision, and the ultimate decision was established as the arithmetic mean of the class label values for each segment. Rounding was conducted according to conventional mathematical rounding rules.

The outcomes of the classifiers constructed using the aforementioned algorithms are presented in (refer to Figure 3).

As evident from the graph above, the classifier based on gradient boosting exhibited the most favorable performance (refer to Figure 4).

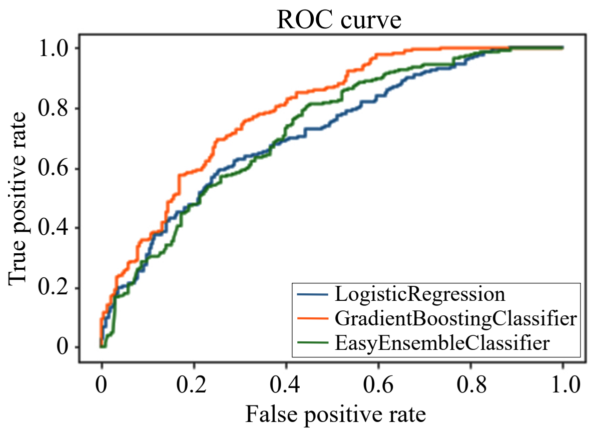

Fig. 3. Classification results based on ROC curves for three classification algorithms: logistic regression, gradient boosting, and EasyEnsembleClassifier (color online)

	precision	recall	f1-score	support
0	0.75	0.66	0.70	243
1	0.69	0.77	0.73	235
accuracy			0.72	478
macro avg	0.72	0.72	0.71	478
weighted avg	0.72	0.72	0.71	478

Fig. 4. Precision-Recall Table for Gradient Boosting Classification

2.1. Utilizing dynamic spectra for classifier construction. The drawback of employing the original rhythmogram recordings is the challenge of achieving temporal synchronization, particularly when segmenting records. Thus, transitioning to a feature space capable of mitigating this limitation appeared logical.

For constructing a classifier rooted in dynamic spectra, the initial signal was partitioned into segments comprising 300 samples each. The spectrum was computed within a window of

100 samples. By incrementing the window by 1 count and computing the spectrum anew within the adjusted window, a spectrogram was generated (refer to Figures 5 and 6).

To construct a classifier based on spectrograms, each spectrogram is transformed into a row vector. As our chosen classification algorithm, we opt for the well-established GradientBoosting-Classifier.

The performance outcome of the constructed classifier surpasses that of the classifier utilizing the original rhythmograms (refer to Figures 7 and 8).

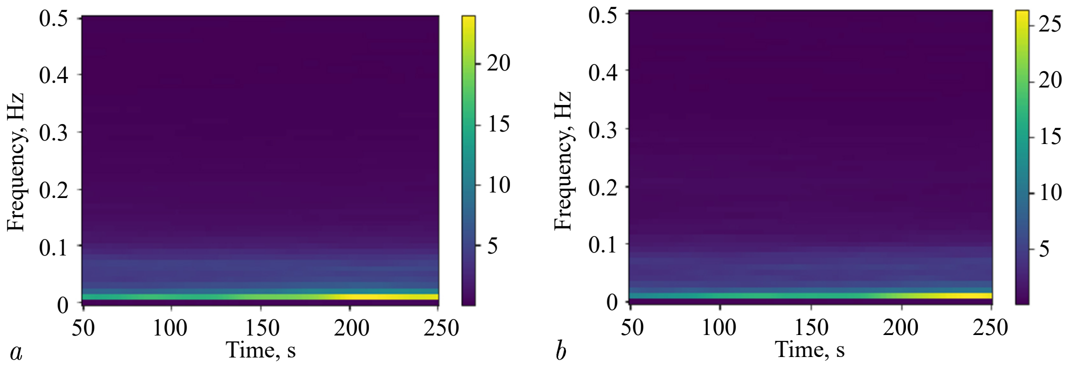

Fig. 5. Averaged spectrogram for training sample records: *a* — belonging to class 0, *b* — belonging to class 1 (color online)

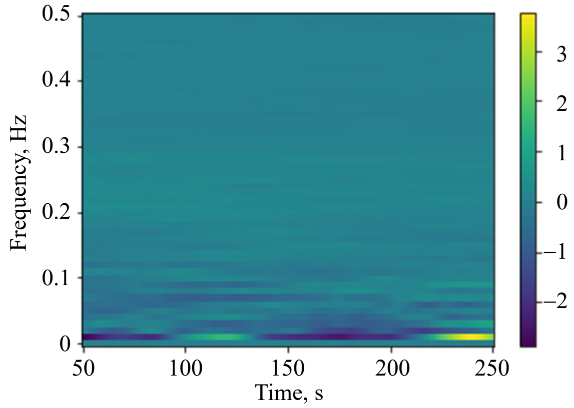

Fig. 6. Difference spectrogram (color online)

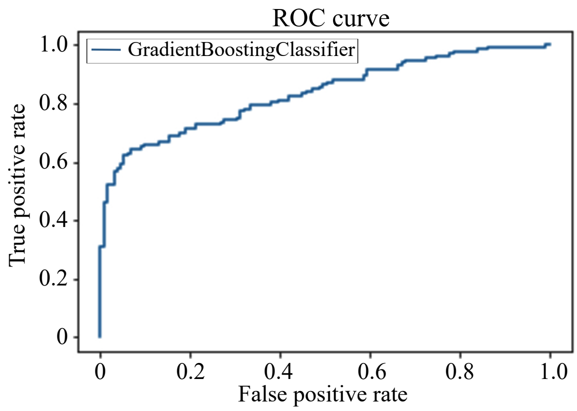

Fig. 7. Classification results based on the ROC curve for the classification algorithm built on spectrograms — gradient boosting

	precision	recall	f1-score	support
0	0.73	0.85	0.78	131
1	0.82	0.68	0.74	132
accuracy			0.76	263
macro avg	0.77	0.76	0.76	263
weighted avg	0.77	0.76	0.76	263

Fig. 8. Precision-Recall Table for Gradient Boosting Classification built on spectrograms

Conclusion

In this paper, we propose a novel method for assessing the level of emotional disadaptation through an individual's cardiorhythmogram. This determination occurs automatically. Information from a portable heart sensor, affixed to an individual, is transmitted via Bluetooth to a mobile device, where the level of emotional disadaptation is ascertained using a pre-trained neural network algorithm. When employing a neural network algorithm, it is recommended to utilize a classifier trained on spectrograms rather than the raw cardiorhythmogram data.

References

1. Miskowiak KW, Johnsen S, Sattler SM, Nielsen S, Kunalan K, Rungby J, Lapperre T, Porsberg CM. Cognitive impairments four months after COVID-19 hospital discharge: Pattern, severity and association with illness variables. European Neuropsychopharmacology. 2021;46:39–48. DOI: 10.1016/j.euroneuro.2021.03.019.
2. Baig AM. Chronic COVID syndrome: Need for an appropriate medical terminology for long-COVID and COVID long-haulers. J. Med. Virol. 2021;93(5):2555–2556. DOI: 10.1002/jmv.26624.
3. Yelin D, Wirtheim E, Vetter P, Kalil AC, Bruchfeld J, Runold M, Guaraldi G, Mussini C, Gudiol C, Pujol M, Bandera A, Scudeller L, Paul M, Kaiser L, Leibovici L. Long-term consequences of COVID-19: research needs. The Lancet Infectious Diseases. 2020;20(10): 1115–1117. DOI: 10.1016/S1473-3099(20)30701-5.
4. Chen C, Haupert SR, Zimmermann L, Shi X, Fritsche LG, Mukherjee B. Global prevalence of post-coronavirus disease 2019 (COVID-19) condition or long COVID: A meta-analysis and systematic review. The Journal of Infectious Diseases. 2022;226(9):1593–1607. DOI: 10.1093/infdis/jiac136.
5. Alkodaymi MS, Omrani OA, Fawzy NA, Shaar BA, Almamlouk R, Riaz M, Obeidat M, Obeidat Y, Gerberi D, Taha RM, Kashour Z, Kashour T, Berbari EF, Alkattan K, Tleyjeh IM. Prevalence of post-acute COVID-19 syndrome symptoms at different follow-up periods: a systematic review and meta-analysis. Clinical Microbiology and Infection. 2022;28(5):657–666. DOI: 10.1016/j.cmi.2022.01.014.
6. Stasenko SV, Kazantsev VB. Covid olfactory dysfunction model. In: 2023 Fifth International Conference Neurotechnologies and Neurointerfaces (CNN). IEEE; 2023. P. 82–85.
7. Stasenko SV. Mean-field model of tripartite synapse with infected glial cells. In: 2023 Fifth International Conference Neurotechnologies and Neurointerfaces (CNN). IEEE; 2023. P. 78–81.
8. Song M. Psychological stress responses to COVID-19 and adaptive strategies in China. World Development. 2020;136:105107. DOI: 10.1016/j.worlddev.2020.105107.
9. Stasenko SV, Hramov AE, Kazantsev VB. Loss of neuron network coherence induced by virus-infected astrocytes: a model study. Scientific Reports. 2023;13(1):6401. DOI: 10.1038/s41598-023-33622-0.
10. Stasenko SV, Kovalchuk AV, Eremin EV, Drugova OV, Zarechnova NV, Tsirkova MM, Permyakov SA, Parin SB, Polevaya SA. Using machine learning algorithms to determine the post-COVID state of a person by their rhythmogram. Sensors. 2023;23(11):5272. DOI: 10.3390/s23115272.
11. Aleksandrovsky YA. Borderline Mental Races Construction. Moscow: Medicina; 2000. 496 p. (in Russian).
12. Parfenov VA, Yakhno NN. General Neurology. Moscow: MIA; 2000. 209 p. (in Russian).

13. Wise TN, Mann LS, Shay L. Alexithymia and the five-factor model of personality. Comprehensive Psychiatry. 1992;33(3):147–151. DOI: 10.1016/0010-440x(92)90023-j.
14. Runova EV, Grigoreva VN, Bakhchina AV, Parin SB, Shishalov IS, Kozhevnikov VV, Nekrasova MM, Karatushina DI, Grigoreva KA, Polevaya SA. Vegetative correlates of conscious representation of emotional stress. Modern Technologies in Medicine. 2013;5(4): 69–76.
15. Grigorieva VN, Tkhostov AS. A method for assessing emotional new human condition. Patent No. 2291720.2007 Russian Federation, IPC A61M21/00: appl. 08.09.2005: publ. 20.01.2007. Assignee: Nizhny Novgorod State Medical Academy of the Federal Agency for Health and Social Development. 23 p. (in Russian).
16. Hosmer DW, Lemeshow S. Applied Logistic Regression. 2nd ed. New York, Chichester: Wiley; 2002. 392 p. DOI: 10.1002/0471722146.
17. Liu X-Y, Wu J, Zhou Z-H. Exploratory undersampling for class-imbalance learning. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics). 2009;39(2):539–550. DOI: 10.1109/TSMCB.2008.2007853.
18. Hastie T, Tibshirani R, Friedman J. Boosting and additive trees. In: The Elements of Statistical Learning. Springer Series in Statistics. New York: Springer; 2009. P. 337–387. DOI: 10.1007/978-0-387-84858-7_10.
19. Schapire RE. Explaining AdaBoost. In: Schölkopf B, Luo Z, Vovk V, editors. Empirical Inference. Berlin, Heidelberg: Springer; 2013. P. 37–52. DOI: 10.1007/978-3-642-41136-6_5.
20. Kramer O. Scikit-learn. In: Machine Learning for Evolution Strategies. Vol. 20 of Studies in Big Data. Cham: Springer; 2016. P. 45–53. DOI: 10.1007/978-3-319-33383-0_5.
21. Lemaître G, Nogueira F, Aridas CK. Imbalanced-learn: A python toolbox to tackle the curse of imbalanced datasets in machine learning. The Journal of Machine Learning Research. 2017;18(1):1–5.