Classification of Brain Activity Patterns Using Machine Learning Based on EEG Data

Marina S. Murtazina, Tatiana V. Avdeenko
Novosibirsk State Technical University, Novosibirsk, Russia

2020 1st International Conference Problems of Informatics, Electronics, and Radio Engineering (PIERE) | 978-1-7281-8990-1/20/$31.00 ©2020 IEEE | DOI: 10.1109/PIERE51041.2020.9314660

Abstract – **The article is devoted to the classification of brain activity patterns using machine learning based on EEG data. The aim of the study is to compare the results of machine learning algorithms depending on the chosen strategy for preparing EEG data and extracting features for the training sample with the example of the identifying fists motor activity task. A brief review of scientific works concerning identification of motor and imaginary activity according to EEG data is presented. An overview of the software for viewing and processing EEG records is also given. The key peculiarities of preliminary preparation of the EEG data are analyzed including removal of artifacts and choice of the channels. The approaches to the selection of a set of features for extraction are analyzed. Description of machine learning tools used to classify patterns of brain activity based on the EEG data is given. During the experiment, the following machine learning methods were studied to classify the EEG data: nearest neighbor method, support vector machine, artificial neural network. Pre-processing of EEG data was performed with use of EDFbrowser. To remove artifacts, a Butterworth banpass filter was used. To extract features, Python software libraries for data processing and analysis were used. WEKA-3-8-4 library of machine learning algorithms was used for training classification models.**

Index Terms – **EEG, preprocessing of data, machine learning, extraction of features.**

I. INTRODUCTION

AT PRESENT, the problems of recognition of the brain electroencephalography (EEG) patterns based on machine learning algorithms is rapidly growing research area concerning the development of applied software, in particular, those used in medicine, education, and gaming industry. EEG is used to identify potentials associated with cognitive and motor events. The main highly specialized areas of cognitive and motor events recognition based on the EEG data are: identification of semantic category of the perceived object [1], recognition of mental tasks [2-3], control of objects using the "brain-computer interface" technology [4]. A prerequisite for building usable pre-trained models with use of various machine learning tools is to ensure the quality of training sample. To form the training sample based on EEG data, it is necessary to determine a set of channels for feature extraction and a set of extracted features. At the same time, feature extraction is carried out from the EEG signals representing a large amount of data with presence of noise artifacts. The choice of channels and extracted features is determined by the specified task of classifying EEG patterns. This paper will consider the classification of the brain activity patterns according to the EEG data using the example of the problem of detecting patterns of fists movement. The purpose of the study is to conduct a comparative analysis of working machine learning algorithms, depending on the chosen strategy of determining the training sample from the EEG data.

The work is organized as follows. Section 1 substantiates the relevance of the research topic. Section 2 formulates the problem statement. Section 3 analyzes the peculiarities of feature extraction from the EEG data, as well as the software for viewing and processing EEG records and the software for machine learning. Section 4 presents the results of the experiment. Section 5 discusses the results obtained. Section 6 draws conclusions about the work done.

II. PROBLEM STATEMENT

There is a lot of research going on in the field of determining certain types of brain activity. Machine learning methods are commonly used to classify the types of brain activity. At the same time, numerous studies gave different results and even conclusions, which can be partially explained by a variety of methods used for preliminary processing of EEG data, the choice of extracted features and machine learning algorithms.

The analysis of EEG patterns of motor or imaginary activity (movements and imaginary movements) is of great importance in medicine, for example, for patients suffering from motor injuries [5], and is also actively studied in the gaming industry [6]. It is believed that imaginary movement leads to the activation of the same areas of the brain as the actual movement [7].

In [8], an approach is proposed for processing and classifying EEG signals during the performance of tasks for imagining movement. From the EEG data, the authors of [8] extract statistical, spatio-temporal and frequency-spatial features, as well as features obtained on the basis of wavelet transform. The work [9] is devoted to the issues of extracting features that make it possible to distinguish between the fist movements performed by the left and right hands using various classification algorithms.

This study is devoted to a comparative analysis of machine learning algorithms used to analyze the EEG. In this case, the study will focus on the problem of preparing a training sample, which includes the selection of channels for making

978-1-7281-8990-1/20/$31.00 ©2020 IEEE

observations and the selection of features for solving the classification problem.

III. THEORY

A. EEG

EEG is now a widely used instrument for recording brain waves. To record the EEG, a number of electrodes are placed on the scalp at the specific points, as well as reference electrodes (usually located on the earlobes) and a ground electrode (usually located on the forehead). In order for scientists to correlate the brain zones and the location of the EEG channels, several international schemes for the location of the electrodes on the human head have been developed. The most popular are the 10-20 and 10-10 systems. These schemes provide for measuring the distance from the bony landmarks of the skull with the subsequent calculation of the intervals between the electrodes in percent. The numbers in the name of the circuit indicate the distance between the electrodes in percent. Electrode names include the first letter of the Latin name for the area where the electrode is placed and a number indicating the side and location of the electrode within that area. The following identifiers are used for brain regions: prefrontal (Fp), frontal (F), temporal (T), parietal (P), occipital (O), central (C). Marker Z is used in the designation of electrodes placed on the median sagittal plane of the skull, marker I is placed on the occipital protuberance, marker A is on the earlobes. Odd numbers of electrodes indicate the left hemisphere of the brain, even numbers indicate the right one. The Fig. 1 shows the arrangement of 64 electrodes in the 10-10 system.

Currently, there is no uniform standard for storing EEG records. Among various storage formats for EEG records, one of the most popular ones among the researchers is the European EDF data format, designed for the exchange and storage of medical time series. EDF data files store numbers corresponding to potentials measured at successive time intervals in each of the EEG channels. In addition, EDF files store additional information in the form of annotations.

The main characteristics of an EEG are frequency and amplitude. Frequency is determined by the number of oscillations per second and is expressed in hertz (Hz). The amplitude is the range of fluctuations in the electric potential on the EEG. Amplitude is measured from the peak of the previous wave to the peak of the next wave in the opposite phase. Amplitude is measured in microvolts (μV).

Traditionally, five main frequency ranges are distinguished in the EEG, denoted by the letters of the Greek alphabet: δ delta (less than 4 Hz), θ theta (from 4 to less than 8 Hz), α alpha (from 8 to less than 13 Hz), β beta (from 13 to less than 30 Hz), γ gamma (more than 30 Hz) [10]. The band range can be subdivided into sub-bands. For example, the beta band is divided into two sub-bands: beta1 (13 Hz to less than 17 Hz) and beta2 (17 Hz to less than 30 Hz). The division into sub-bands may differ in different sources, both in terms of the boundaries of the sub-ranges, and in their number. The fluctuations released in the general electrical activity of the brain are called rhythms. Rhythms differ from each other in wavelength, amplitude, shape, frequency and functional meaning. For example, the alpha rhythm is the rhythm of the brain in the frequency range from 8 to 13 Hz with an amplitude of 40-70 μV. In the EEG of a healthy adult in a state of calm wakefulness, the alpha rhythm is most pronounced in the occipital regions. In addition to the alpha rhythm itself, several more rhythms are observed in the same frequency range, which appear in other areas of the brain and have a different waveform (mu rhythm, kappa rhythm, tau rhythm).

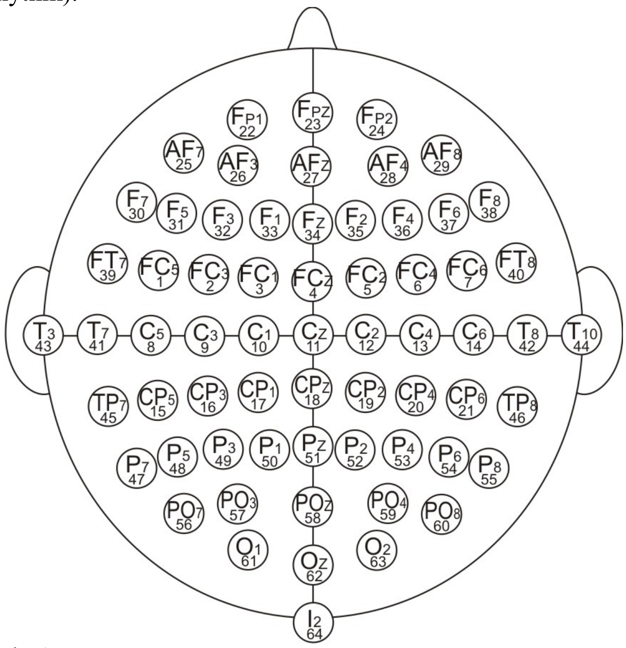

Fig. 1. EEG electrode diagram according the 10-10 system.

The most important frequency ranges in the context of the problem of identifying motor activity from the EEG data are the following ranges: mu (9-13 Hz), alpha (8-12 Hz), beta (13-30 Hz) and gamma (> 30 Hz) [11]. It should be noted that the frequency ranges and the boundaries of the frequency ranges differ in the studies of different authors. Thus, in [12], it is proposed to use two frequency ranges mu (7-13 Hz) and beta (13-30 Hz) to study the patterns of motor activity.

The motor cortex is responsible for motor activity. Therefore, when identifying motor events, attention is focused on the EEG data obtained from electrodes located above the motor cortex. In [13], electrode nodes and frequencies were investigated to determine the direction of hand movement. According to the results of the parameters analysis obtained from the brain functional network (BFN), it was found that the most sensitive appeared to be the EEG components the delta, theta and gamma1 frequencies from the F4, F8, C3, Cz, C4, CP4, T3 and T4 electrodes. In [14], an approach to decoding the observed or imaginary hand movements based on the data obtained from the FC3, FC4, C5, C3, C1, C2, C4, C6, CP3, CP4 electrodes is proposed using the power spectral density in the mu and beta ranges.

B. *Software for viewing, processing and analyzing EEG data*

For the purposes of viewing, processing and analyzing EEG data, many software tools have been created. The most popular open source programs in the scientific community are: EEGLab, FieldTrip, BrainStorm [15]. The listed programs are an interactive Matlab toolbox. EEGLab is released under the BSD Free License, FieldTrip is under the GNU Public License, Brainstorm is under the GNU Public License. Brainstorm is currently available in two forms: an open source Matlab application and a standalone Java executable. Using Matlab and Java makes Brainstorm a cross-platform application. The presence of a Java implementation makes the latter tool the easiest to use for the end user.

EDFBrowser is one of the most popular programs for working with EEG data in EDF format. EDFBrowser is a free multi-platform program for viewing and processing physiological data such as EEG data.

Many Python libraries for working with EEG data are also being actively developed. Among them, the most popular are MNE-Python and PyEEG. MNE-Python is an open source software package that supports a set of methods for preprocessing the EEG data, source localization, statistical analysis, and assessment of functional connectivity between distributed brain regions [16]. PyEEG is an open source Python library for feature extraction from the EEG data [17].

C. *Machine learning for solving classification problems based on the EEG data*

Classification is the most common machine learning task, the essence of which is to build models that assign an object of interest to one of several known classes. The classical scheme of machine learning organization for solving classification problems based on EEG data is shown in Fig. 2.

The raw EEG recording data is a multidimensional time series with discrete time. The number of EEG channels determines the dimension of each point in the time series. The number of points in the time series depends on the duration of the EEG recording and the frequency of its sampling. For example, for an EEG recording 1 minute long at a sampling rate of 160 Hz, the number of points in the time series will be 9 600 points. This raw data may contain artifacts that need to be filtered out.

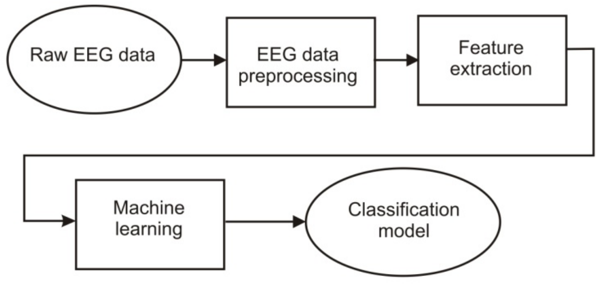

Fig. 2. Organization of machine learning based on the EEG data.

Artifacts are unwanted signals arising from the registration of electrical activity that is not of cerebral origin. Artifacts can be divided into artifacts of physiological origin (such as blinking eyes, muscle activity, heartbeat) and non-physiological origin (such as environmental noise, experimental errors) [18-19]. The main sources of nonphysiological artifacts caused by environmental noise are external electrical noise (for example, from power line wires) and internal electrical faults in the recording system due to recording electrodes [20]. Bandpass filtering is one effective approach to removing artifacts. Cardiac artifacts can occur when electrodes are placed on or near a blood vessel. The frequency of cardiac artifacts is about 1.2 Hz [18]. Blinking artifacts are represented by signals with a frequency below 4 Hz [21]. The noise in the frequency spectrum by muscle artifacts starts at about 20 Hz. With an increase in the frequency value, an increase in noise is observed [22]. Non-physiological artifacts caused by electrical noise are observed at frequencies around 50 Hz or 60 Hz [23].

The choice of channels through which feature extraction is performed depends on the ultimate goals of the study, which are quite diverse. The main goals when choosing data acquisition channels are to improve performance, to reduce general dimension, and to determine the area of the brain that generates a certain type of activity [24].

The extracted features should reflect characteristic signal parameters that can be used to identify brain states associated with a specific task. As EEG features, indicators from different areas related to time series analysis can be used: power spectral density from signal processing, fractal measurements from computational geometry, entropy from information theory, and so on. A common set of features that are often extracted from the EEG recording for the analyzed frequency ranges are: average, median, maximum and minimum values for the frequency range [25]. The extraction of features is performed after segmentation of the EEG data [26], during which periods of time (epochs) corresponding to the investigated brain activities are highlighted in the EEG recording.

After forming a dataset for machine learning, it is necessary to choose learning algorithm and to test the resulting model.

D. *Software for machine learning*

For machine learning from the EEG data, there are specialized libraries for the classification of biosignals, as well as the software that is not intended specifically for the field of EEG analysis, is actively used. A striking example of the first group of the software tools is the pyRiemann machine learning library [27] which provides a high-level interface for classifying EEG biosignals using Riemann geometry.

From the second group of software, the WEKA package is popular among the researchers [28]. WEKA is an open source java application that is a collection of machine learning algorithms for data mining tasks. This software is widely used in various application areas: from financial data analysis to biomedical one. For WEKA machine learning tool, the input is usually provided in a file with the *.ARFF (Attribute-Relation File Format) extension. The *.ARFF format is specially designed for use with Weka machine learning software. The *.ARFF format file consists of two sections. The first section contains the name of the relationship, a list of attributes (columns in the data) and their types. The second section is the rows of data instances.

WEKA is an open source java application that is a collection of machine learning algorithms for data mining tasks. This software is widely used in various subject areas: from financial data analysis to biomedical one. For WEKA machine learning tool, the input is usually provided as a file with the *.ARFF (Attribute-Relation File Format) extension. The *.ARFF format is specially designed for use with Weka machine learning software. The *.ARFF format file consists of two sections. The first section contains the name of the relationship, a list of attributes (columns in the data) and their types. The second section contains rows of data instances.

IV. Experimental Results And Discussion

For the experiment, data from the “EEG Motor Movement /Imagery” dataset, created by the developers of the BCI2000 instrumental system [29], were used. The dataset is placed in the PhysioNet publicly available data warehouse for medical research [30].

The signals from the electrodes located above the motor cortex were selected for analysis. The selected channels are FC2, FC4, FC6, C2, C4, C6, CP2, CP4, CP6, FC5, FC3, FC1, C5, C3, C1, CP5, CP3 and CP1. EDFbrowser was used to extract a set of specified channels from the original record. To remove artifacts, a Butterworth bandpass filter was applied.

To extract features from the EEG records, we chose a set consisting of the following characteristics: maximum value, minimum value, average value, average of the absolute values of the first signal differences, the average of the absolute values of the second signal differences, the average of the absolute values of the first differences of the normalized signal, average from the absolute values of the second differences of the normalized signal, standard deviation.

The above features were extracted from the EEG data record for the first subject from the “EEG Motor Movement/Imagery” dataset. Before feature extraction, the EEG data were segmented. The segment length is determined according to the following rule: for recording the state of "rest" it is used 4.2 seconds, for recording the movement "clenching the fist" it is used 4.1 seconds. The sampling rate is 160 Hz. Thus, the total of 30 segments were extracted: 15 records of EEG data for the state "rest", 8 records for the state "clenching the left fist", 7 records for the state "clenching the right hand fist". The Python library was used to automate the feature extraction process.

From the total set of extracted features, three data sets were formed, describing the same set of EEGs. The first data set includes the following features: maximum value, minimum value, average of absolute values of the first signal differences, average of absolute values of the second signal differences, standard deviation. The second set of features includes: maximum value, minimum value, average value, average of the absolute values of the first signal differences. The third set of features includes: maximum value, minimum value, average value, average of the absolute values of the first differences of the normalized signal, the average of the absolute values of the second differences of the normalized signal, standard deviation. The datasets have been aligned with the *.ARFF format requirements.

Machine learning was performed on each dataset using the WEKA tool. The following algorithms were used for training: nearest neighbor method, support vector machine, Bayesian network, multilayer perceptron. During classification, the standard settings of the WEKA machine learning tool were used. The model training results are shown in Fig. 3-5

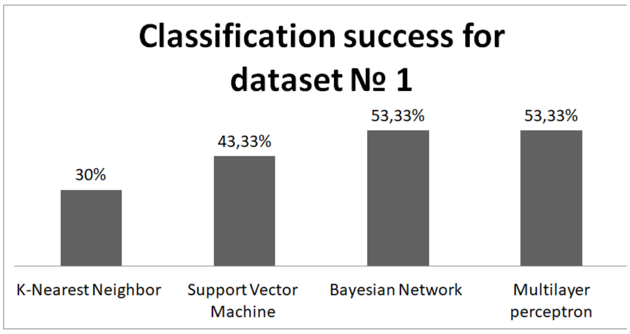

Fig. 3 Comparison of the classification success with the selected algorithms for dataset № 1.

In Fig. 3-5 the results of classification success (percentage of correctly classified objects) for three sets of features, and for four different classification methods are presented. Analyzing the results obtained, one can conclude that the set of features, as well as the classification method, has a significant impact on the classification results.

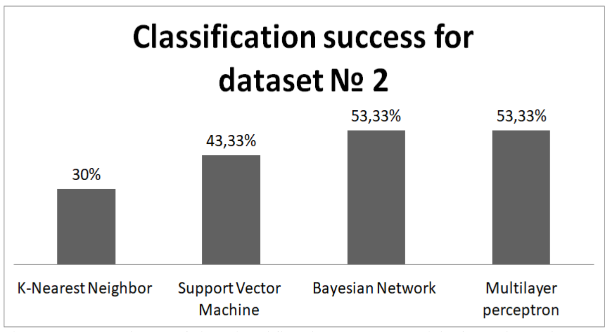

Fig. 4. Comparison of the classification success with the selected algorithms for dataset № 2.

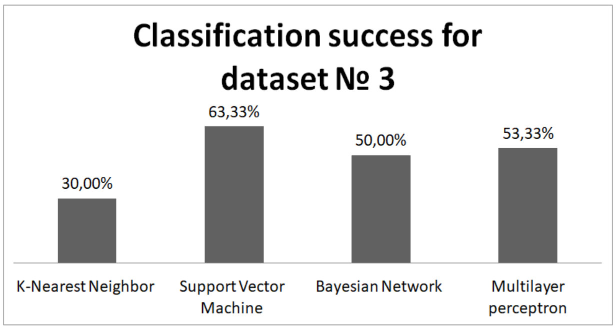

Fig. 5 Comparison of the classification success with the selected algorithms for dataset № 3.

As for the set of features, the third set is the best. The features included in it turned out to be the most informative in the aggregate. Even better results can be obtained in the future by applying the methods of automatic selection of the most informative features, for example, the method of principal components.

Among the classification methods, the Bayesian network turned out to be the best for solving this problem. Support vector machines and multilayer perceptron give approximately equal results. The nearest neighbours method turned out to be unsatisfactory, which is most likely due to the localization of the obtained optimum. It should be noted here that we used general classification methods, without taking into account the specifics of the problem being solved. A more thorough fit and adaptation of classification methods to the analyzed data can significantly increase the success of the classification.

A more detailed analysis of the results revealed that classification errors occurred when distinguishing between the state 1 "rest" and one of the states 2 "clenching the fist of the left hand" or 3 "clenching the fist of the right hand". There were practically no errors in distinguishing between the states 2 and 3. This may be due to the fact that there is a certain transition between the state of rest and the states of motion, which was not taken into account in the preprocessing of the signal. This observation can be taken into account when planning the EEG segmentation procedure, and can significantly affect the classification success.

VI. CONCLUSION

The study allows us to state the prospects of applying machine learning methods to EEG data obtained when performing various types of activities. It is advisable to associate further research with the improvement of the EEG segmentation procedure, with the development and application of a systematic method for selecting informative features, with the adaptation of machine learning methods to solving specific problems.

ACKNOWLEDGEMENTS

The research is supported by Ministry of Science and Higher Education of Russian Federation (project No. FSUN-2020-0009).

REFERENCES

[1] I. Simanova et al., “Identifying object categories from event-related EEG: toward decoding of conceptual representations,” *PloS one*, vol. 5(12), e14465, Dec 2010, doi: 10.1371/journal.pone.0014465.

[2] W. Qiang and O. Sourina, “Real-time mental arithmetic task recognition from EEG signals,” *IEEE transactions on neural systems and rehabilitation engineering: a publication of the IEEE Engineering in Medicine and Biology Society*, vol. 21(2), Jan. 2013, pp. 225–32, doi: doi:10.1109/TNSRE.2012.2236576.

[3] N. N. Neshov et al., “Classification of Mental Tasks from EEG Signals Using Spectral Analysis, PCA and SVM,” *Cybernetics and Information Technologies*, vol. 18(1), pp. 81–92, Mar. 2018, doi: https://doi.org/10.2478/cait-2018-0007.

[4] M. Z. Ilyas et al., “Classification of EEG signals for brain-computer interface applications: Performance comparison,” *Int. Conf. on Robotics, Automation and Sciences (ICORAS)*, Ayer Keroh, 2016, pp. 1–4, doi: 10.1109/ICORAS.2016.7872610.

[5] M.Z. Bai, et al., “Filtering techniques for channel selection in motor imagery EEG applications: a survey”, *Artif Intell Rev*, vol. 53, pp. 1207–1232, Feb. 2020, doi: https://doi.org/10.1007/s10462-019-09694-8.

[6] N. Padfield et al., “EEG-Based Brain-Computer Interfaces Using Motor-Imagery: Techniques and Challenges,” *Sensors (Basel, Switzerland)*, vol. 19(6), pp. 1423, Mar. 2019, doi: doi:10.3390/s19061423.

[7] Th. Mulder, “Motor imagery and action observation: cognitive tools for rehabilitation,” *Journal of Neural Transmission*, vol. 114(10), pp. 1265–78, Oct. 2007, doi:10.1007/s00702-007-0763-z.

[8] S. R. Sreeja et al., “Motor Imagery EEG Signal Processing and Classification Using Machine Learning Approach,” *Jordanian Journal of Computers and Information Technology*, vol. 4(2), pp. 80–93, Aug. 2018, doi: 10.5455/jjcit.71-1512555333.

[9] M. H. Alomari et al., “Automated Classification of L/R Hand Movement EEG Signals using Advanced Feature Extraction and Machine Learning,” ArXiv , abs/1312.2877 (2013): n. pag.

[10] P. Campisi and D. L. Rocca, “Brain waves for automatic biometric-based user recognition,” in *IEEE Transactions on Information Forensics and Security*, vol. 9(5), pp. 782-800, May 2014, doi: 10.1109/TIFS.2014.2308640.

[11] R. Scherer and & C. Vidaurre, “Motor Imagery based Brain-Computer Interfaces,” In P. Diez (Ed.), Smart Wheelchairs and Brain-Computer Interfaces: Mobile Assistive Technologies. Academic Press, 2018, pp. 171-195. doi: https://doi.org/10.1016/B978-0-12-812892-3.00008-X.

[12] F. Stefano Carlos et al., “EEG sensorimotor rhythms' variation and functional connectivity measures during motor imagery: linear relations and classification approaches,” *PeerJ*, vol. 5, e3983, Nov. 2017, doi:10.7717/peerj.3983.

[13] Li. Ting et al., “Decoding Voluntary Movement of Single Hand Based on Analysis of Brain Connectivity by Using EEG Signals,” *Frontiers in Human Neuroscience*, vol. 12, pp. 381, Nov. 2018. doi:10.3389/fnhum.2018.00381.

[14] A. Korik et al., “Decoding Imagined 3D Hand Movement Trajectories From EEG: Evidence to Support the Use of Mu, Beta, and Low Gamma Oscillations*,” Frontiers in Neuroscience*, vol. 12, pp. 130, Mar. 2018, doi:10.3389/fnins.2018.00130.

[15] R. Oostenveld et al., “FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data,” Computational Intelligence and Neuroscience, vol. 2011, 9 pages, 2011, doi:10.1155/2011/156869.

[16] A. Gramfort et al.,“MEG and EEG data analysis with MNE-Python,” *Frontiers in Neuroscience*, vol. 7, p. 267, Dec. 26, 2013, doi:10.3389/fnins.2013.00267.

[17] F. S. Bao et al., "PyEEG: an open source Python module for EEG/MEG feature extraction," *Computational Intelligence and Neuroscience*, vol. 2011, pp. 406391, 2011, doi: 10.1155/2011/406391.

[18] X. Jiang et al., "Removal of Artifacts from EEG Signals: A Review," *Sensors* (Basel, Switzerland), vol. 19(5), pp. 987, Feb. 2019, doi:10.3390/s19050987.

[19] J. A. Urigüen and G.-Z. Begoña, "EEG artifact removal-state-of-the-art and guidelines," *Journal of Neural Engineering*, vol. 12(3), pp.: 031001, Apr. 2015, doi:10.1088/1741-2560/12/3/031001.

[20] E. S. Papathanasiou et al., "Knitting artifact," *Journal of Neurology*, neurosurgery, and psychiatry, vol. 74(11), pp. 1501, 2003, doi:10.1136/jnnp.74.11.1501.

[21] M. Shahbakhti et al., "Elimination of Blink from EEG by Adaptive Filtering without Using Artifact Reference," *2013 4th Int. Con. on Intelligent Systems*, *Modelling and Simulation*, Bangkok, 2013, pp. 190–194, doi: 10.1109/ISMS.2013.106.

[22] S. D. Muthukumaraswamy, "High-frequency brain activity and muscle artifacts in MEG/EEG: a review and recommendations," *Frontiers in Human Neuroscience*, vol. 7, pp. 138, Apr. 2013, doi:10.3389/fnhum.2013.00138.

[23] P. Schembri et al. "Detection of Electroencephalography Artefacts using Low Fidelity Equipment," PhyCS, 2017, doi:10.5220/0006398500650075.

[24] T. Alotaiby et al., "A review of channel selection algorithms for EEG signal processing," EURASIP J. Adv. Signal Process, art. 66 (21 p.), Aug. 2015. doi: https://doi.org/10.1186/s13634-015-0251-9/

[25] V. Podgorelec, "Analyzing EEG Signals with Machine Learning for Diagnosing Alzheimer's Disease", *Elektron Elektrotech*, vol. 18(8), pp. 61–64, Oct. 2012.

[26] A. Procházka et al., "Multi-channel EEG signal segmentation and feature extraction," *2010 IEEE 14th Int. Conf. on Intelligent Engineering Systems*, 2010, pp. 317–320.

[27] M. Mousavi et al., "Spatio-temporal analysis of error-related brain activity in active and passive brain–computer interfaces," *Brain-Computer Interfaces*, vol. 6(4), pp. 118–127, 2019.

[28] L. Piatek et al., "Eye state classification from electroencephalography recordings using machine learning algorithms," *Digit Med*, vol. 4, pp. 84–95, 2018.

[29] G. Schalk et al., "BCI2000: A General-Purpose Brain-Computer Interface (BCI) System," *IEEE Transactions on Biomedical Engineering*, vol. 51(6), pp. 1034–1043, 2004.

[30] A. Goldberger et al., "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation* [Online], vol. 101 (23), pp. e215–e220, 2000.

Marina Sh. Murtazina was born in Chita City, in 1984. She graduated from Chita State University (now Transbaikal State University, TransSU, Chita, Russia), Department of Applied Informatics. In 2012 she received her Candidate of Science Degree in Philosophical anthropology, philosophy of culture, from TransSU. From 2004 to 2016 she worked in TransSU. Since 2016 she is an Associate Professor of the Automated Control Systems Department and a Senior Researcher of Scientific and Educational Center "Intelligent Information Technologies in Business", Novosibirsk State Technical University. She is an author of more than 70 scientific papers and 3 textbooks. Her research interests include Intelligent Information Systems, Decision Support Systems, Knowledge Engineering, Agile Software Engineering.

Tatiana V. Avdeenko was born in Novosibirsk City, in 1963. She graduated from Novosibirsk Electrical Engineering Institute (now Novosibirsk State Technical University – NSTU, Novosibirsk, Russia), Department of Applied Mathematics. In 1991 she received her Candidate of Science Degree in Control of Technical Systems, from NSTU. In 2004 she received her Doctor of Science Degree in Mathematical Modelling, Numerical Methods and Software Packages, also from NSTU. From 2005 to 2016 she was a Head of Economic Informatics Department of NSTU. At present she is a Full Professor of the Applied Mathematics and Computer Science Department and a Chief Researcher of Scientific and Educational Center "Intelligent Information Technologies in Business". She is an author of more than 200 scientific papers including 1 monograph and 7 textbooks. Her research interests include Data Mining, Knowledge Engineering, Artificial Intelligence, Information and Decision-Support Systems.