Supervised machine learning algorithms for brain signal classification

Ihab Abdulrahman Satam[a], *Róbert* Szabolcsi[b]

[a] Obuda University, Doctoral School on Safety and Security Sciences, Budapest, Hungary;
Northern Technical University, Al-hawija Technical Institute, Electronic Techniques Department, Hawija, Republic of Iraq,
e-mail: ihab.satam@uni-obuda.hu, **corresponding author,**
ORCID iD: https://orcid.org/0000-0002-9749-0944

[b] Obuda University, Doctoral School on Safety and Security Sciences, Budapest, Hungary,
e-mail: szabolcsi.robert@bgk.uni-obuda.hu,
ORCID iD: https://orcid.org/0000-0002-2494-3746

https://doi.org/10.5937/vojtehg72-48620

FIELD: biomedical engineering
ARTICLE TYPE: original scientific paper

Abstract:

Introduction/purpose: The brain wave application is widespread in recent years, especially in the applications that aid the impaired people suffered from amputation or paralysis. The objective of this research is to assess how well different supervised machine learning algorithms classify brain signals, with an emphasis on improving the precision and effectiveness of brain-computer interface applications.

Method: In this work, brain signal data was analyzed using a number of well-known supervised learning models, such as Support Vector Machines (SVM) and Neural Networks (NN). The data set was taken from a previous study. Twenty five participants imagined moving their right arm (elbow and wrist) while the brain signals were recorded during that process. The dataset was prepared for the analysis by the application of meticulous pre-processing and feature extraction procedures. Then the resulting data were subjected to classification.

Results: The study highlights how crucial feature selection and model modification are for maximizing classification results. Supervised machine learning methods have great potential for classifying brain signals, particularly SVM and NN.

Conclusion: The use of SVM and NN has the potential to completely transform the creation of cutting-edge brain-computer interfaces. The integration of these models with real-time data should be investigated in future studies.

Keywords: supervised machine learning, EEG, brain signals, classifications, feature extraction.

ACKNOWLEDGMENT: The author is grateful for the support of the Obuda University for their suport in this research.

Introduction

The field of electroencephalography (EEG) signal processing lies at the crossroads of neuroscience and machine learning, with far-reaching implications for comprehending and interpreting complicated cerebral activities. This work describes a game-changing breakthrough in the categorization of EEG data accomplished by the methodical use of supervised machine learning techniques. The research expands on a dataset from a prior study, in which the categorization accuracy of EEG signals was 65%. This study effectively increased classification accuracy to 95% by utilizing sophisticated algorithms such as Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and Optimized Ensemble techniques. This large improvement in accuracy not only contributes significantly to the study of neural signal analysis, but it also opens up new pathways for practical applications in medical diagnostics, neurotherapeutic methods, and the creation of complex brain-computer interfaces. The improved capacity to identify EEG data with such high precision is critical in interpreting the complicated patterns of brain activity, with implications for understanding and treating neurological illnesses, improving cognitive capacities, and building adaptive neurotechnology.

The brain signals are oscillating electrical impulses occurring in the brain tissue; these impulses have different frequencies depending on the mental state or the activity. Slow waves are related to deep sleep while fast waves are associated with active thoughts and speech. The brain is connected to the rest of the organs through nerves; these nerves are just like telephone wires that connect homes together all around the world. When a human being wants to move an organ, let us say the hands, the brain sends signals to the hands to inform the muscles to contract. In fact, the nerves send a lot of electrical impulses called action potentials to different muscles in the hands allowing the movement in extreme precision.

A Brain-Computer interface (BCI) or Brain-Machine interface (BMI) is a device that is designed to read or decode a signal from the brain to directly control external devices such as Prosthetics, Wheelchair, or a robot. The BCI system acquires signals from the brain and, with the help of the signal processing techniques, translates them into control commands that provide real-time feedback to the user.

The chain process of the common BCI (Figure 1) system has several stages starting with:

- Amplification and digitizing of the analogue signal collected from the electrodes using analogue-to-digital conversion,

- Signal processing of the signal (removing noise and artifacts),
- Feature Extraction, and
- Classification.

The BCI has several types depending on the signal recording methods: Invasive, Semi-Invasive, and Non-Invasive. The Invasive and Semi-Invasive methods provide more accurate signals, higher bandwidth, better spatial resolution, and less artifacts than the Non-Invasive one. The Non-Invasive method is generally easier in signal recordings and does not require surgical intervention as the previous two methods.

Machine Learning (ML) algorithms are programs that can learn the hidden patterns from the data, predict the output and improve the performance. There are different types of algorithms used in ML as shown in Figure 2. This figure shows three types of algorithms: Supervised, Unsupervised and Reinforcement learning.

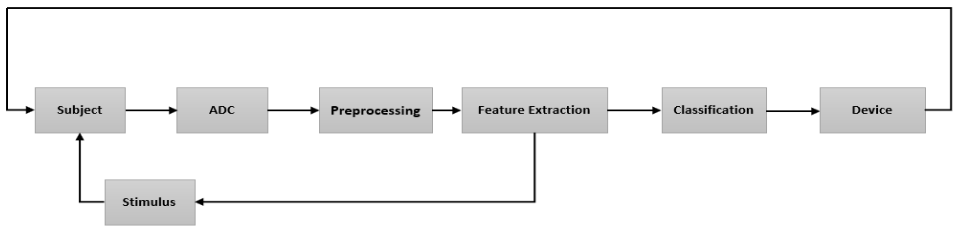

Figure 1 – BCI system components

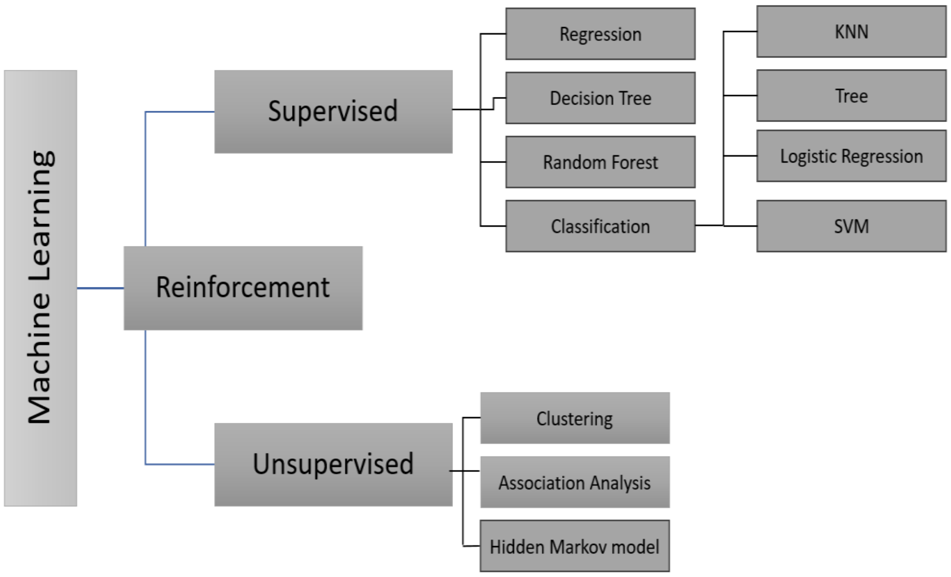

Figure 2 – ML types

In Supervised learning, as indicated by the name, the machine needs external supervision to learn. The Supervised learning models are trained using labeled datasets. Once training and processing are done, the model is tested by providing sample test data to check whether it predicts the correct output or not. Supervised learning is divided into: Classification and Regression. Unsupervised learning is the type where the machine does not need external supervision to learn from the data. The unsupervised model can be obtained using the unlabeled dataset that is not classified, nor categorized and the algorithm needs to act on that data without any supervision. The types of unsupervised algorithms are Clustering and Association. Reinforcement learning is a type where the agent interacts with the environment by performing actions and learn with the help of feedback. There are several studies regarding using ML algorithms for the classification of brain data. Amin et al (2017) proposed a pattern recognition approach. They used four types of classifiers: K-nearest neighbors (KNN), Support Vector Machine (SVM), Multi-layer Perceptron (MLP), and Naïve Bayes (NB). For low frequencies 0-3.91 Hz, the accuracy achieved using SVM for A5 approximate coefficient was 99.11%. For detailed coefficients D5, two algorithms were used: SVM and KNN with the accuracy of 98.57% and 98.39%, respectively. For higher frequencies, 3.91-7.81 Hz, the other algorithms were used for A5 and D5 coefficients. MLP(A5-D5) with 97.11-89.63%, respectively, and NB(A5-D5) with 91.6-81.07%, respectively. De Brito Guerra et al (2023) developed a machine learning model based on the Random Forest algorithm to classify EEG signals from subjects performing real and imagery motor activities. In another work by Ramírez-Arias et al (2022), they used 9 machine learning algorithms to classify a signal related with a real motor movement. They indicate that medium Artificial Neural Network was the best algorithm with the area under the curve of 0.9998 and losses with 0.0147. Huang et al (2020) presented a classification methodology using sparse representation and Fast Compression Residual Convolutional Neural Networks (FCRes-CNNs). Their work achieved the accuracy of 98.82%. Wen et al (2021) proposed a deep network model for autonomous learning and the classification of EEG signals. The model can self-adapt to classify EEG signals with different frequencies. Despite the artificial design feature, the extraction method was not able to obtain stable classification results; the model had a better classification accuracy with short length signals. Paez-Amaro et al (2022) used four algorithms: Common Spatial Patterns (CSP) combined with Linear Discriminant Analysis (LDA), Deep Neural Network (DNN), Convolutional Neural Network (CNN), and finally

Riemannian Minimum Distance to Mean (RMDM). The mean accuracy for these algorithms were 8%, 66%, 60% and 80%, respectively.

In this study, six ML algorithms were used to increase the accuracy of the former study from Ma et al (2020). The algorithms used were: Optimal SVM, Fine SVM, Optimal KNN, Optimal Decision Tree, Optimal NN, and optimized Ensemble.

Methodology

Dataset

The data for this work was made available by Ma et al (2020) and each trial had been approved by the ethics board at the Institute of Automation, Chinese Academy of Sciences. A total of 25 healthy, right-handed subjects were polled. (19 males, 6 females). The participant had no Mi-based BCI knowledge. It is important to mention that all the experiments were approved by the Ethical Committee of the Institute of Automation, Chinese Academy of Sciences.

Method of recording signals

The data signals were selected from Ma et al (2020), as mentioned in the previous section. In order to completely comprehend the scenario, it is critical to describe the way by which the signals were recorded. The subjects reclined in a comfortable chair, their hands naturally resting on their legs, and looked at the screen from one meter away, Figure 3(a). Every trial lasted 8 seconds and began with a white circle in the middle of the screen for two seconds, as illustrated in Figure 4.

Then, a red circle flashed for one second as a signal to assist people focus on the coming goal. Before the desired reaction was required, the "Hand" or "Elbow" cue was presented for 4s. The participants were instructed to visualize doing the needed activity with their entire bodies, rather than simply their eyes, during this time. The participants were told to relax their limbs and think about whatever they wished. EMGs were recorded from the right hands and forearms of the subjects, Figure 3(b), to ensure they were not acting on their own initiative. (The EMG signals were eliminated during the EEG preprocessing.

Finally, the "Break" for 1s was enough to put an end to the 8s experiment and relax. During the interval, patients were instructed to relax and reduce ocular and muscular activity.

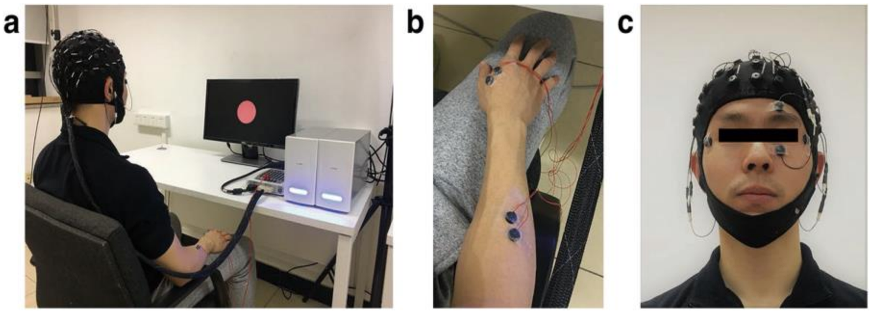

Figure 3 – Data recording

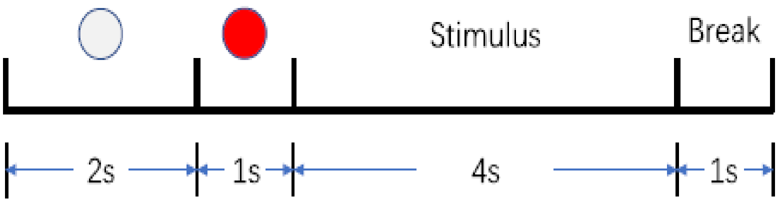

Figure 4 – Time period of one trial

The motor-related processes alpha (8-13 Hz) and beta (14-30 Hz) are associated with event-related synchronization and desynchronization (ERS/ERD). When the actual action is performed or anticipated, ERD manifests as a decrease in a certain frequency component associated with an increase in brain activity. What makes any enhanced frequency sensitive, or ERS, is an increase in a certain frequency component. Because it is connected to the suppression of brain activity, it can sometimes be detected even when no actual action is being done or imagined.

Researchers have recently uncovered a strong correlation between a specific sensory brain area and the ability to envisage moving certain parts of the body (as shown in Figure 5). In the picture, the dark blue region in the middle of the brain is in charge of directing the motion of the limbs. The pale cyan area is in charge of guiding hand movement.

After everything is said and done, the region above the ears is in charge of moving things like the cheekbones and the lips. ERD in the dominant brain and ERS in the nondominant hemisphere can be caused by motor imagery. Figure 6 depicts the spectrum strength of the brain frequency ranges (indicated in red).

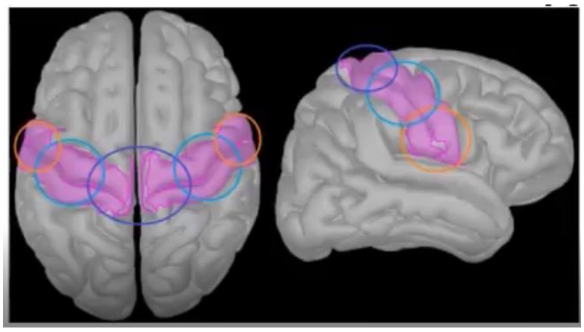

Figure 5 – Regions of motor imagery

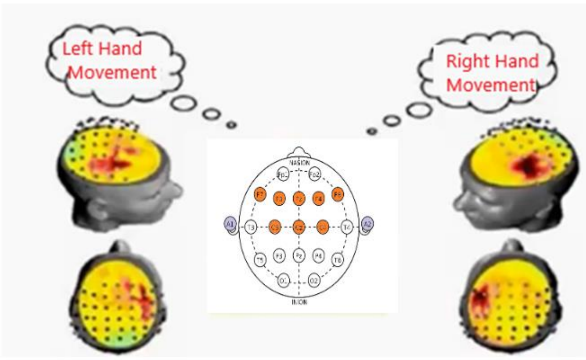

Figure – 6 BCI System for motor imagery signal recording

Using a Neuroscan SynAmps2 amplifier and a 64-channel electrode cover, the conventional 10/20 system was utilized to acquire EEG data at a sampling rate of 1000 Hz. For the electroencephalogram (EEG) recordings, the left mastoid was employed as a reference. Throughout the trial, the electrode impedances were maintained below 10 k ohm. The acquired data was cleaned up in MATLAB (R2015a) using the EEGLAB toolbox (v14.1.1_b). We used a common average reference (CAR) in the early phases of processing. A low-pass filter with a frequency of 40 hertz and a high-pass filter with a frequency of 0.1 hertz were installed. To save processing expenses, the input was down sampled to 200Hz. The EEG was cleaned of anomalies associated to the eyes and muscles using automatic artifact removal (AAR). The data set was preprocessed in [7] and became ready for feature extraction. The wavelet transform is a non-stationary time-scale analytic method that may be used with EEG data. It is extremely useful to be able to evaluate non-stationary signals and break them down into discrete frequency components over a wide range of timeframes. WT allows researchers to condense complicated biological signals comprised of several time-varying data sets into a digestible collection of diagnostic variables (Amin et al, 2017).

The continuous and discrete Wavelet transform formula are both given in equations (1) and (2).

$$WT_x(a,\tau) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t)\psi * \frac{(t-\tau)}{a} dt \quad (1)$$

Continuous Wavelet Transform

where a represents scale displacement, τ represents time displacement, and ψ is a wavelet basis function, including Hear, db.Series, Coiflet, etc.

$$WT_x(j,k) = \int x(t)\psi_{j,k}^{*}(t)dt \quad (2)$$

Discrete Wavelet Transform

The DWT (Discrete Wavelet Transform) is used to discretize scale and displacement by limiting the end of the wavelet basis function (a,). Figure 7 displays the DWT breakdown of the EEG signal x(n). Convolution is a two-function multiplication method that uses the low-pass or high-pass filter coefficients and is then processed using own sampling. To down sample, the sample signal must be cut in half (reduced). Wavelet signals are classified into two types: approximation signals and detail signals. An approximation is a signal produced by the convolution of the original signal with a low-pass filter, whereas a detail is produced by the convolution of the original signal with a high-pass filter. Figure 7 shows that each output generates a detailed signal D and an approximation signal A, with the more recent one acting as the input for the phase that follows. The main frequency component of the EEG data dictates how many layers the wavelet decomposes.

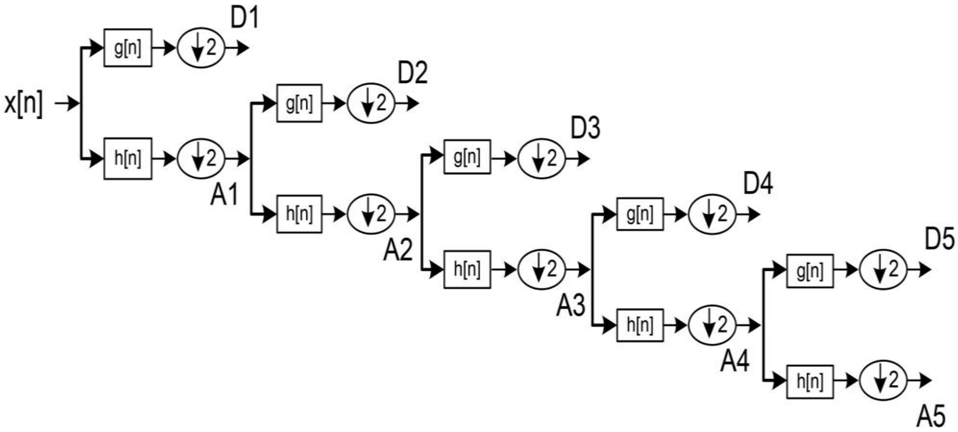

Figure 7 – Wavelet decomposition

The formula of WT and the filter h, is a low pass and can be formulated in the formulation as follows:

$$H(z)\,H(z-1)\;+\;H(-z)\,H(-z-1)\;=\;1 \qquad (3)$$

In the above formula, $H(z)$ is used to represent the h, z-transform filter and the complement transformation of this high-pass filter is expressed as:

$$G(z)\;=\;zH(-z-1). \qquad (4)$$

According to the Section above, DWT is used to evaluate the spectrum components of EEG data. The EEG signal analysis relies heavily on WT, specifically the careful selection of a wavelet and the optimal number of breakdown stages. The number of thresholds is calculated based on the primary frequency component of the EEG data. In order to classify signals, the levels are chosen such that the wavelet coefficients maintain a strong connection between the various parts of the signals and the requisite frequencies. The analysis was performed using five distinct degrees of decomposition. Therefore, the EEG data is segmented into D1-D5 details and a final method, A5.

Multiple wavelet varieties are typically tested to find the most effective combination for a particular application. As a result of its Daubechies wavelet feature, second order (db2) filtering is more adept at detecting variations in the input signal. Therefore, wavelet coefficients were generated using db2 for this study.

For the Daubechies wavelet of the second order (db2), the band frequencies are as follows, with a sampling frequency of 256 Hz: D1 (64-128 Hz); D2 (32-64 Hz); D3 (16-32 Hz); D4 (8-16 Hz); D5 (4-8 Hz); and A5 (2-4 Hz). (0 - 4 Hz). To determine discrete wavelet values, MATLAB is used. Because even the most effective classifier will fail with a badly selected input feature, this is a crucial factor in the design of artificial neural networks based on pattern categorization. Determining the wavelet discontinuous coefficient provides a representation of the signal's energy across time and frequency. For this reason, the discontinuous wavelet coefficient calculated from the EEG signal of each record serves as the feature vector used to characterize the signal. The size of the recovered feature vector is reduced by using statistics on top of the collection of wavelet coefficients. The temporal frequency distribution of the signals under study is represented by the statistical characteristic listed below:

- Means and standard deviation value,
- Variance,
- Skewness,
- Kurtosis, and
- Root Mean Square.

In this paper, the data needed for the right arm is collected from C3 channel as explained in Section 3. In order to retrieve the features for the EEG data prior to classification, a code is developed in MATLAB for this wavelet.

Means and Standard Deviation Value

The definition of the mean is very simple as it is the sum of all the signals divided by the number of the signals.

$$\mu = \frac{1}{N}\sum_{i=0}^{N-1} X_i \tag{5}$$

The expression |X_i-μ| indicates the difference between the deviation of the sample and their mean. The average deviation can be found by the sum of all the derivative of the sample signals and dividing by the total number of samples. The standard deviation is similar but the average is done by power instead of amplitude as shown in equation (6).

$$\sigma = \sqrt{\frac{1}{N-1}\sum_{i=0}^{N-1}(x_i - \mu)^2} \tag{6}$$

Variance

It is the variability measure. In order to determine the variance, the average cubed departure from the mean is used as the denominator. The extent of dispersion in a data collection can be better understood by examining its variance. The variance from the mean increases as data spreads out.

$$\sigma^2 = \frac{1}{N-1}\sum_{i=0}^{N-1}(X(i) - \mu)^2 \tag{7}$$

Skewness

Skewness is a statistical measure of the degree to which a signal deviates from its mean value. To compute it, divide the cubed standard deviation by the cubed mean variation.

$$\gamma = \frac{1}{(N-1)\sigma^3}\sum_{n=0}^{N-1}(x_n - \mu)^3 \tag{8}$$

Kurtosis

It is the Kurtosis of the signal that determines its Peakedness. More peaks in the waveform correspond to a greater kurtosis number.

$$K = \frac{\frac{1}{N-1}\sum_{i=0}^{N-1}(x_i^4)}{(\frac{1}{N-1}\sum_{i=0}^{N-1}(x_i^2))^2} \tag{9}$$

Root Mean Square

It is a quantitative representation of the signal's intensity. The signal's magnitude is determined using the root-mean-square formula. The strength is represented by the range. The root-mean-square deviation provides a measure of the variability in the system's response to external factors.

$$R.M.S = \sqrt{\frac{1}{N-1}\sum_{i=0}^{N-1} x_i^2} \quad (10)$$

Machine learning algorithms

Machine learning algorithms are a critical component of the fast-evolving science of artificial intelligence. These algorithms allow computers to learn from data and make predictions or choices without the need of human instructions. Here is a quick rundown of some of the most common types of machine learning algorithms:

1. Supervised learning: In this sort of learning, algorithms are trained using labeled training data—that is, data that has already been assigned the right answer. Based on this data, the algorithm predicts things and is adjusted if it gets it wrong. For continuous outputs, common examples include logistic regression or support vector machines for classification methods, and for categorical outputs, linear regression.
2. Unsupervised learning: In this type of learning, data is sent to the algorithm without clear instructions on how to handle it. Finding some organization in the data through exploration is the aim. Clustering methods like K-means or hierarchical clustering, as well as dimensionality reduction algorithms like Principal Component Analysis (PCA), are examples of common unsupervised learning algorithms.
3. Semi-Supervised learning: This method falls in between unsupervised and supervised learning. Here, both labeled and unlabeled data are used to train the algorithm. When obtaining a properly labeled dataset is costly or time-consuming, this approach might be helpful.
4. Reinforcement learning: In this method, an algorithm gains decision-making skills by acting in a way that advances its objective in the environment. Rather than being informed directly of its errors, it learns from the results of its actions. This method is frequently applied in gaming, navigation, and robotics.

5. Deep learning: A branch of machine learning that makes use of multi-layered neural networks, or "deep networks," to assess different aspects of data. Deep learning is very effective for voice and picture recognition applications.

Every kind of machine learning algorithms has advantages and disadvantages, and it may be used to many kinds of jobs. The type of data and the particular needs of the work are major factors in selecting an algorithm. In this paper, the supervised machine learning algorithms were applied to the data for classification. The main idea of Supervised ML is mapping between the input output of that data. In order to accomplish this task, the algorithm receives training data. Training data consists of input and output pairs. Inputs are multidimensional vectors that represent relevant information about the signal states (which in our case are the brain signal states) or activities to be decoded. Typically, raw data is used to create features which are then manually optimized through feature engineering to identify the most promising or relevant ones. Training involves learning the mapping between features and desired outcomes. The response variable, also known as the dependent variable, refers to the output of interest linked with these traits, such as brain state or behavior. During the training phase, the model learns to map the input characteristics to target variables by optimizing model parameters, which can be achieved through minimizing the cost function. The model performance is indicated by the loss or error estimated by the cost function. There are several algorithms that have the ability to minimize the cost function, but the most common one is gradient descent. An acceptable machine algorithm model consists of training and testing sets. The testing set should never be evaluated by the algorithm during the training phase. It should accurately depict the model's actual setting. The testing results should be as good as the training sets. Figure 8 represents the machine learning architecture.

The algorithms used were Support Vector Machine and Neural Network.

Support Vector Machine

Although it may also be used for regression, Support Vector Machine (SVM) is a strong and adaptable supervised machine learning technique that is mostly utilized for classification problems. Given that it can handle both linear and non-linear data, many difficult classification tasks benefit greatly from its use. SVM tends to find the most adequate dividing boundary (Linear or Hyperplane). SVM locates the hyperplane with the highest margin for a two-class classification issue, which indicates that it is the farthest from the closest data points in each class. Support vectors, which are these closest points, are essential for establishing the hyperplane (Ahmed et al, 2022; Al-Hamadani, 2023).

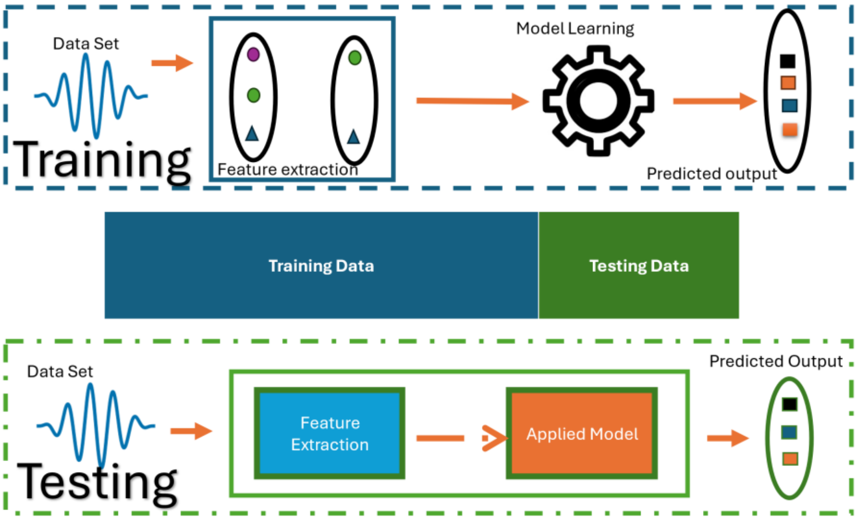

Figure 8 – Architecture of machine learning

Neural Network

A neural network that is intended to be optimized or enhanced over time is referred to as an "optimizable neural network". A neural network's optimization often focuses on enhancing its performance in certain tasks, such as predictive modeling, language processing, or image recognition (Ortega-Fernandez et al, 2024) Below are some key concepts related to optimizable neural networks:

1- Learning algorithm: An optimizable neural network's learning algorithm is its most important component. This approach, which is frequently a variation of gradient descent, modifies the network's weights in response to input. The objective is to reduce the deviation, sometimes referred to as the cost or loss, between the network's predicted and actual results.

2- Backpropagation: Backpropagation is a common optimization technique in neural networks, particularly in deep learning models. It is a technique for quickly calculating the gradients of the loss function in relation to the network weights. The weights are then updated using this knowledge to lessen the loss.

3- Hyperparameters: These are the configurations or settings that control the neural network's general behavior but are left alone while it learns. Learning rate, batch size, and the number of network layers are a few examples. An essential step in neural network optimization is fine-tuning of these hyperparameters.
4- Regularization and Overfitting: Overfitting is the process by which a network gets excessively specialized to the training set and underperforms when exposed to fresh, untested data. Through the use of regularization techniques like dropout, weight decay, or early halting, overfitting is prevented, improving the network's generalizability and real-world performance.
5- Data Preprocessing: The performance of the network can be greatly impacted by the way data is prepared and displayed. Normalization, standardization, and augmentation are some of the techniques that can improve the effectiveness and efficiency of network training (Al-Aziz et al, 2021).
6- Transfer Learning: Using a pre-trained model on a sizable dataset and then honing it for a particular job is a common method used to enhance neural networks. This strategy can result in substantial performance gains, particularly when there is a dearth of data available for the particular job.
7- Evolutionary Algorithms: To optimize neural networks, certain cutting-edge techniques include evolutionary algorithms. By producing a population of networks, picking the top-performing ones, and then using those networks to build a new generation of networks, these techniques iteratively enhance the performance of the network (Kuptsov & Stankevich, 2024).
8- Hardware Optimization: Lastly, employing GPUs (Graphics Processing Units) or TPUs (Tensor Processing Units) for quicker processing—a critical step in the effective training of big networks—the performance of neural networks may also be enhanced at the hardware level.

In the recent years, the development of artificial intelligence presented an efficient opportunity in the enhancement of the lives of disabled individuals. However, ethical considerations must be pointed out. To begin, we must recognize the importance of ethics in AI research. When designing technology to aid disabled people, it is critical to examine the possible implications for their privacy, autonomy, and well-being. Failure to effectively address ethical issues can cause harm to the very people these technologies are intended to serve. Thus, ethical monitoring and thought are essential in AI research. Furthermore, privacy and data security considerations are critical when dealing with sensitive personal information about disabled people. Data anonymization, encryption, and rigorous access restrictions must be applied to ensure the confidentiality of personal data. Furthermore, gaining informed consent and guaranteeing individuals' ownership over their data are critical ethical issues for protecting autonomy and privacy rights. Addressing the ethical concerns in AI research targeted at enhancing the everyday activities of disabled people is critical to ensuring the appropriate and useful deployment of technology. Researchers can uphold ethical values and maximize the positive impact of AI technologies on the lives of disabled people by recognizing potential risks such as bias, privacy concerns, and dependency, as well as implementing mitigation strategies such as fairness principles, data security measures, and stakeholder engagement. Ethical monitoring and reflection must be integrated into AI research to promote inclusion, autonomy, and dignity for all humans.

Results

The study's findings are provided in this section. The methods and goals described in the preceding sections are followed in the organization and detail of the results. To make sure that the data was fully understood, a variety of analytical approaches were used. The purpose of this part is to lay the foundation for the debate and interpretation of the study findings by offering an objective and factual summary of the findings. As the analysis and ramifications of these findings will be fully addressed in the discussion section that follows, it is crucial to highlight that the data is provided without prejudice or subjective interpretation.

The feature extraction explained in the previous sections aims to find the best features of the data that can be used in classification later. Figure 9 shows the feature extracted using equations 5 – 10.

The features extracted help the classification algorithms to train in order to find a pattern that can predict the outcome for each trial. The data was labeled as 1 and 2 which represents if the subject imagined moving his elbow or his wrist. Figures 10 and (11-12) show the SVM and NN algorithms respectively for classifications.

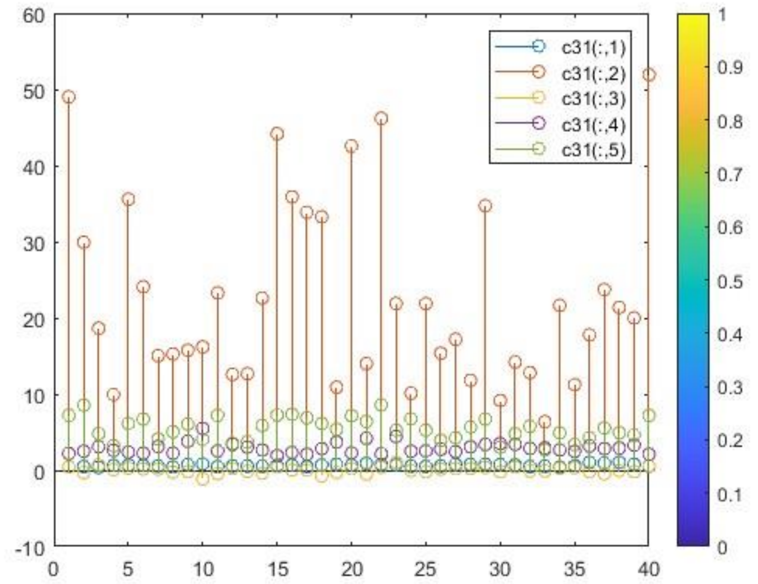

Figure 9 – Features extracted

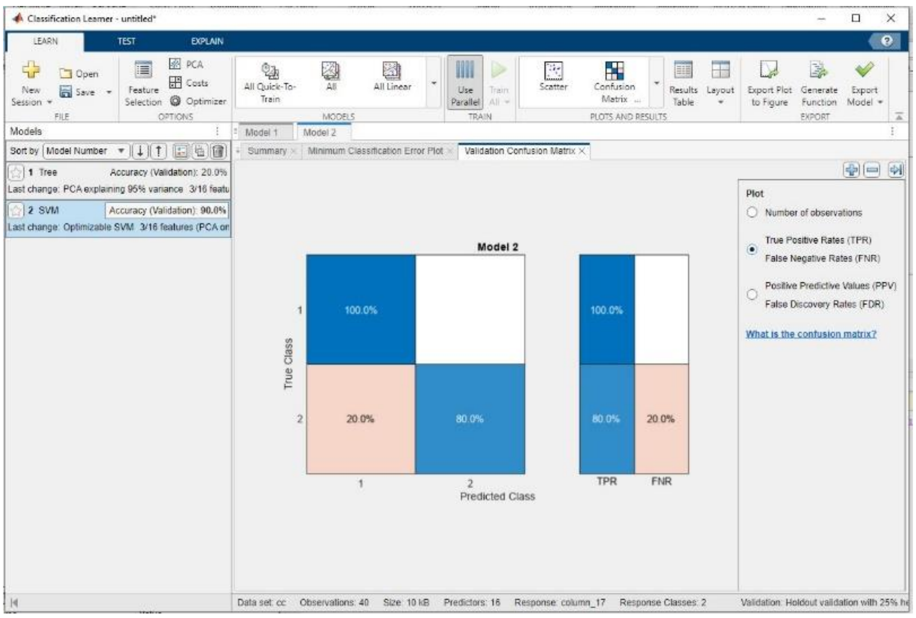

Figure 10 – SVM

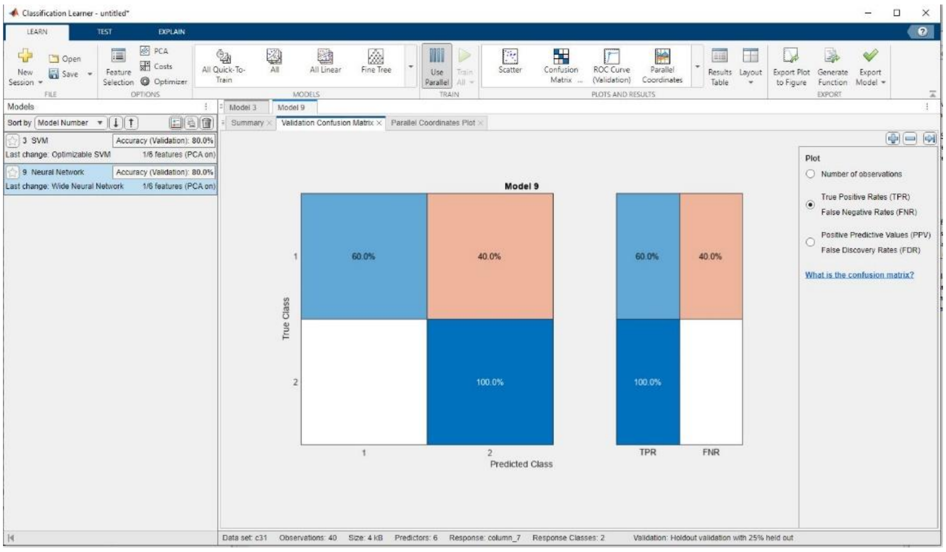

Figure 11 – Neural network 1

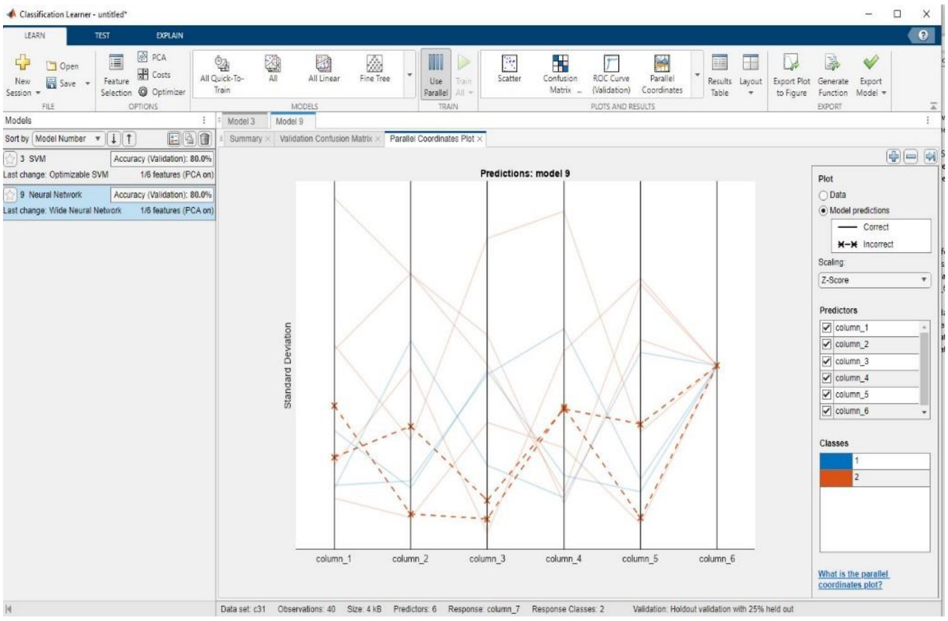

Figure 12 – Neural Network 2

The figures show the results for only one subject. Table 1 present all subjects' results after applying the same signal processing, feature extraction and classifications.

Table 1 – Classification accuracy

No	Accuracy	
	SVM	NN
1	90.00%	80.00%
2	85.00%	84.00%
3	85.00%	82.00%
4	80.00%	80.00%
5	84.00%	81.00%
6	82.00%	82.50%
7	75.00%	84.00%
8	80.00%	83.00%
9	79.00%	84.00%
10	76.00%	87.00%
11	85.00%	80.00%
12	80.00%	80.00%
13	77.00%	80.00%
14	85.00%	81.00%
15	73.00%	86.00%
16	78.00%	90.00%
17	88.00%	85.00%
18	84.00%	81.00%
19	86.00%	83.00%
20	91.00%	81.00%
21	96.00%	81.00%
22	98.00%	81.00%
23	85.00%	85.00%
24	96.00%	82.00%
25	82.00%	81.00%

Figure 13 shows the difference between the SVM, and NN results compared to the results shown in Ma et al (2020). The improvement made in this study shows a good step up in classifying the signals.

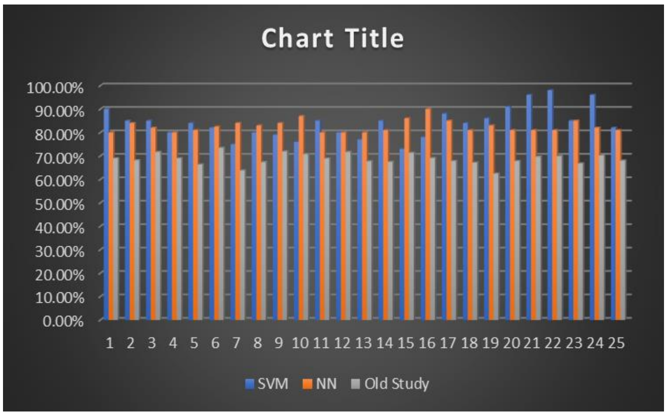

Figure 13 – Comparison of the results

Conclusion

The applications of brain waves has become widespread in recent decades. The usage of these programs enhances the daily activities of disabled individuals. Achieving high levels of accuracy is essential. The non-invasive method of signal recording is the most commonly used in this sort of application. In this procedure, EEG electrodes are used to record signals from various areas of the brain while the brain is functioning in response to a stimulus. This work improved a prior study's classification accuracy of brain signals. Different algorithms are employed to analyze signals, extract features, and classify them. The findings demonstrate how the algorithms improved the classification accuracy of the prior study, indicating that the work was on the right path. Achieving this accuracy percentage leads to the development of new approaches for future accuracy increase. There are several sorts of algorithms that may be used for classification, but in our study, we only employed two since they produced the best results. The findings obtained were based on the type of characteristics extracted.

References

Ahmed, A.H., Al-Hamadani, M.N.A. & Satam, I.A. 2022. Prediction of COVID-19 disease severity using machine learning techniques. *Bulletin of Electrical Engineering and Informatics*, 11(2), pp.1069-1074. Available at: https://doi.org/10.11591/eei.v11i2.3272.

Al-Hamadani, M.N. 2023. Classification and analysis of the MNIST dataset using PCA and SVM algorithms. *Vojnotehnički glasnik/Military Technical Courier*, 71(2), pp. 221-238. Available at: https://doi.org/10.5937/vojtehg71-42689.

Amin, H.U., Mumtaz, W., Subhani, A.R., Saad, M.N.M. & Malik, A.S. 2017. Classification of EEG Signals Based on Pattern Recognition Approach. *Frontiers in Computational Neuroscience*, 11, art.number:103, pp.1-12. Available at: https://doi.org/10.3389/fncom.2017.00103.

Al-Aziz, A., Muntakim, A.-H. & Ahmed, K. 2021. No Regular Behavior Pattern in Neural Network Execution – A Matlab Experience. *International Journal of Computer Application*, 174(19), pp.39-46. Available at: https://doi.org/10.5120/ijca2021921087.

de Brito Guerra, T.C., Nóbrega, T., Morya, E., de M. Martins, A. & de Sousa, V.A. 2023. Electroencephalography Signal Analysis for Human Activities Classification: A Solution Based on Machine Learning and Motor Imagery. *Sensors*, 23(9), art.number:4277. Available at: https://doi.org/10.3390/s23094277.

Huang, J.-S., Li, Y., Chen, B.-Q., Lin, C. & Yao, B. 2020. An Intelligent EEG Classification Methodology Based on Sparse Representation Enhanced Deep Learning Networks. *Frontiers in Neuroscience*, 14, art.number:808. Available at: https://doi.org/10.3389/fnins.2020.00808.

Kuptsov, P.V. & Stankevich, N.V. 2024. Modeling of the Hodgkin–Huxley neural oscillators dynamics using an artificial neural network. *Izvestiya VUZ Applied Nonlinear Dynamics,* 32(1), pp.72-95. Available at: https://doi.org/10.18500/0869-6632-003079.

Ma, X., Qiu, S. & He, H. 2020. Multi-channel EEG recording during motor imagery of different joints from the same limb. *Scientific Data*, 7, art.number:191, pp.1-9. Available at: https://doi.org/10.1038/s41597-020-0535-2.

Ortega-Fernandez, I., Sestelo, M. & Villanueva, N.M. 2024. Explainable generalized additive neural networks with independent neural network training. *Statistics and Computing,* 34, art.number:6. Available at: https://doi.org/10.1007/s11222-023-10320-5.

Páez-Amaro, R.T., Moreno-Barbosa, E., Hernández-López, J.M., Zepeda-Fernández, C.H., Rebolledo-Herrera, L.F. & de Celis-Alonso, B. 2022. EEG motor imagery classification using machine learning techniques. *Revista Mexicana de Fisica*, 68(4), art.number:041102. Available at: https://doi.org/10.31349/RevMexFis.68.041102.

Ramírez-Arias, F.J., García-Guerrero, E.E., Tlelo-Cuautle, E., Colores-Vargas, J.M., García-Canseco, E., López-Bonilla, O.R., Galindo-Aldana, G.M. & Inzunza-González, E. 2022. Evaluation of Machine Learning Algorithms for Classification of EEG Signals. *Technologies,* 10(4), art.number:79. Available at: https://doi.org/10.3390/technologies10040079.

Wen, T., Du, Y., Pan, T., Huang, C. & Zhang, Z. 2021. A Deep Learning-Based Classification Method for Different Frequency EEG Data. *Computational and Mathematical Methods in Medicine*, 2021, art.ID:1972662. Available at: https://doi.org/10.1155/2021/1972662.

Algoritmos de aprendizaje automático supervisados para la clasificación de señales cerebrales

Ihab Abdulrahman Satam[ab], *Róbert* Szabolcsi[a]

[a] Universidad de Obuda, Escuela de Doctorado en Ciencias de la Seguridad y la Protección, Budapest, Hungría

[b] Universidad Técnica del Norte, Instituto Técnico Al-hawija, Departamento de Técnicas Electrónicas, Hawija, República del Iraq

CAMPO: Bioingeniería
TIPO DEL ARTÍCULO: artículo científico original

Resumen:

Introducción/objetivo: La aplicación de las ondas cerebrales está muy generalizada en los últimos años, especialmente en las aplicaciones que ayudan a las personas discapacitadas que han sufrido alguna amputación o parálisis. El objetivo de esta investigación es evaluar qué tan bien diferentes algoritmos de aprendizaje automático supervisado clasifican las señales cerebrales, con énfasis en mejorar la precisión y efectividad de las aplicaciones de interfaz cerebro-computadora.

Métodos: En este trabajo, los datos de las señales cerebrales se analizaron utilizando varios modelos bien conocidos de aprendizaje supervisado, como las máquinas de vectores de soporte (SVM) y las redes neuronales (NN). El conjunto de datos se tomó de un estudio anterior. Veinticinco participantes imaginaron mover su brazo derecho (codo y muñeca) mientras se registraban las señales cerebrales durante ese proceso. El conjunto de datos se preparó para el análisis mediante la aplicación de meticulosos procedimientos de preprocesamiento y extracción de características. Luego los datos resultantes fueron sometidos a clasificación.

Resultados: El estudio destaca cuán cruciales son la selección de características y la modificación del modelo para maximizar los resultados de clasificación. Los métodos de aprendizaje automático supervisados tienen un gran potencial para clasificar señales cerebrales, particularmente SVM y NN.

Conclusión: El uso de SVM y NN tiene el potencial de transformar completamente la creación de interfaces cerebro-computadora de vanguardia. La integración de estos modelos con datos en tiempo real debería investigarse en estudios futuros.

Palabras claves: aprendizaje automático supervisado, EEG, señales cerebrales, clasificaciones, extracción de características.

Контролируемые алгоритмы машинного обучения в классификации сигналов мозга

Ихаб Абдулрахман Сатам[аб], **корреспондент,** *Роберт* Саболчи[а] [а] Обудский университет, докторантура по технике безопасности, г. Будапешт, Венгрия

[б] Северный технический университет, Технический институт Аль-Хавиджа, департамент электронной техники, г. Хавиджа, Республика Ирак

РУБРИКА ГРНТИ: 34.57.00 Биоинженерия
ВИД СТАТЬИ: оригинальная научная статья

Резюме:

Введение/цель: Применение BrainWave в последние время сильно возросло, особенно в области приложений, которые помогают людям с ограниченными возможностями, перенесшим ампутацию или паралич. Цель данного исследования − оценить, насколько точно различные контролируемые алгоритмы машинного обучения классифицируют сигналы мозга с акцентом на повышение точности и эффективности при взаимодействии мозга и компьютера.

Методы: В данной статье проанализированы данные о сигналах мозга с использованием ряда хорошо известных моделей контролируемого обучения, таких как методы опорных векторов (МОВ) и нейронные сети (НС). Данные заимствованы из предыдущего исследования. Двадцать пять респондентов представляли, как двигают правой рукой (локоть и запястье), а одновременно записывались сигналы мозга. Набор данных был подготовлен для анализа путем тщательной предварительной обработки и выявления признаков. Затем полученные данные были подвергнуты классификации.

Результаты: Исследование подчеркивает значимость выбора признаков и модификации модели для получения наилучших результатов классификации. Методы контролируемого машинного обучения имеют большой потенциал в классификации сигналов мозга, особенно методы МОВ и НС.

Выводы: Применение *методов* МОВ *и* НС *может полностью изменить процесс* разработки *современных интерфейсов «мозг-компьютер».* А внедрение *этих моделей с данными*, полученными в *реальном времени, должно стать предметом дальнейших исследований.*

Ключевые слова: контролируемое машинное обучение, ЭЭГ, сигналы мозга, классификация, выявление признаков.

Надгледани алгоритми машинског учења за класификацију можданих сигнала

Ихаб Абдулрахман Сатам[аб], **аутор за преписку**, *Роберт* Саболчи[а] [а] Универзитет Обуда, Школа докторских безбедносних студија, Будимпешта, Мађарска

[б] Северни технички универзитет, Технички институт Ал-Хавиџа, Одељење електронске технике, Хавиџа, Република Ирак

ОБЛАСТ: биоинжењеринг
КАТЕГОРИЈА (ТИП) ЧЛАНКА: оригинални научни рад

Сажетак:

Увод/циљ: Мождани таласи имају све чешћу примену, наричито у области апликација за помоћ особама са ампутацијом или парализом. Циљ овог истраживања јесте процена успешности класификовања можданих сигнала различитих надгледаних алгоритама машинског учења, са тежиштем на побољшању прецизности и ефикасности интерфејса између мозга и рачунара.

Методе: У раду су анализирани подаци можданих сигнала помоћу неколико познатих надгледаних модела учења, попут метода потпорних вектора (SVM) и неуралних мрежа (NN). Подаци су узети из претходне студије. Двадесет пет испитаника је замишљало да покреће десну руку (лакат и ручни зглоб) док су истовремено снимани мождани сигнали. Скуп података је припремљен за анализу коришћењем детаљних поступака претходне обраде и екстракције карактеристика.

Резултати: Студија наглашава суштински значај селекције карактеристика и модификације модела за добијање што бољих резултата класификације. Надгледане методе машинског учења имају велики потенцијал за класификовање можданих сигнала, нарочито методе SVM и NN.

Закључак: Коришћење метода SVM и NN има потенцијал да потпуно трансформише креирање најсавременијих интерфејсова између мозга и рачунара. Интеграција ових модела са подацима добијеним у реалном времену треба да буде предмет будућих истраживања.

Кључне речи: надгледано машинско учење, ЕЕГ, мождани сигнали, класификације, екстракција карактеристика.

Paper received on: 09.01.2024.
Manuscript corrections submitted on: 07.06.2024.
Paper accepted for publishing on: 08.06.2024.