IJARCCE

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 12, December 2021

DOI: 10.17148/IJARCCE.2021.101252

EEG Analysis for BCI and Epilepsy Detection

Chetana R[1], Mahantesh K[2], Channabasava[3]

Research scholar, Department of ECE, SJB Institute of technology, Bengaluru, India[1]

Assosciate Professor, Department of ECE, SJB Institute of technology, Bengaluru, India[2]

BE, Department of ECE, SJB Institute of technology, Bengaluru, India[3]

Abstract: A system capable of translating the logical activity into messages or commands for reactive applications also commonly called as Brain-Computer Interface (BCI). The project is divided in three phases. First Phase deals with Dataset collection, Binary class dataset analysis and Methodology (includes pre-processing of data, classification model, efficiency evaluation). Second phase discusses about complex dataset analysis (cognitive control, Prediction on finger flexion), Advanced algorithms for feature extraction and classification. Third phase discusses about Fine tuning of Algorithms, develop user Interface, Deploy ML model to interface. Different techniques like PCA (principal component analysis), Statistical Approach, Discrete Wavelet Transform (DWT) is used in extracting feature which is the basic techniques used in binary class datasets. Comparison and authentication between fundamental Machine Learning (ML) procedure using K- Nearest Neighbors KNN is done to give better prediction accuracy than the complicated Machine learning procedures like (Support Vector Machine (SVM), Artificial Neural Network (ANN), or Deep Neural Network (DNN). Results obtained for average efficiency for epilepsy detection using statistical and wavelet as features and SVM as classifier is 98%. Considering eye state detection dataset, obtained an average efficiency of 55%. Similarly for cognitive dataset an average of 95% efficiency was obtained by using PCA as a feature and KNN as classifier. In finger flexion dataset the aim was to find out the correlation between finger moments, by using the N best channel frequency pairs and applying the linear regression for the dataset, an average of 0.38 correlation was obtained. Thus, different method and classifiers have obtained fair predictions and efficiency was used. After training of ML models, the user interface was built using HTML and CSS and the trained models were connected to the interface by means of flask framework. The interface gives the user, freedom to select different combination of datasets and features and predict the class accordingly. The interface is also facilitated with visualization of the input test signal which uses chart for plotting graphs.

Keywords: BCI, Feature extraction, Classifiers, User Interface

I. INTRODUCTION

BCI, commonly known as brain machine interface (BMI), enables communication to happen between outer atmosphere and individual through software and hardware communications system, without indulging perimetric nervous and muscular system, and considering only regulatory indications in the form of signals whose origin is the Brain in the form of EEG brain action [6]. BCI is also a communication system comprising of hardware and software that allows straight way of communicating with human's brain and control computers or peripheral devices [1]. Primary issues investigated in the project is analysis of the brainwaves, understanding the characteristics of various waveforms generated in brain as per the different frequency ranging from 3Hz and above (delta, theta, alpha, beta, gamma) due to the properties of signals being fired from the neurons for different tasks.

Clear understanding the flow of the machine learning models and different techniques for fine tuning of the trained model for brain signals is being discussed[7]. Identified the suitable feature extraction techniques for appropriate datasets. Evaluation of efficient classifier models using features extracted from the brain waves based on the signal analysis[2]. Developed a user interface for choosing different datasets, feature extraction techniques and classifiers combination and connect our ML model to interface and make predictions of the different classes of the datasets for the chosen combination[3]. Seizure detection using EEG, EEG Eye State, cognitive control, Prediction on finger flexion datasets for our analysis of signals used considered datasets. Simple methods involving machine learning is implemented to segregate and classify brain signals using Electroencephalography (EEG) data.

II. DATASETS CONSIDERED

A. Epilepsy detection dataset

The EEG dataset is of 500 files and 5 different folders of 100 each files are categorized according to patients was considered for the analysis which is available online from University Bonn, Germany. Each single channel EEG signal recording of 23.6s is considered with 173.61 Hz sampling frequency to generate 4097 data points. The EEG brain waves ranges from 0.5 Hz to 85 Hz. After sampling, to obtain 4097 datapoints sampling in time-series is done. The recording for

IJARCCE

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 12, December 2021

DOI: 10.17148/IJARCCE.2021.101252

the healthy and epileptic patients both are to be considered. The EEG recording on several occasion of time during sampling is the data point.

Once sampling is done the entire range of 4097 data points are considered and split in bins of 23 numbers, where all bin possesses 1 second of brain activity at different point in time as 178 datapoints which is considered for experimentation. The dataset is reconstructed, dividing, and shuffled. This comprises of 11500 samples considered in all the 5 folders. So now we have 23 bins x 500 files = 11500 samples of information row wise, and the columns of 1 second are represented by 178 datapoints of which label Y is the column in the end. which takes the value from 1 to 5 representing 5 different classes. Last column represents the dependent variable Y and X1, X2, …, X178 represents the independent variables
Class 5- Eyes open of healthy volunteers EEG signal data is considered.
Class 4- Eyes closed of healthy volunteers EEG signal data is considered.
Class 3- Healthy brain's EEG signal data is considered.
Class 2- Existence of brain tumour location spot, EEG data is considered
Class 1- Epileptic seizure segments, EEG signals are considered

B. Eye state detection dataset
Emotive EEG Neuroheadset is used for collecting the continuous EEG measurement[9]. The recording is done for 117 seconds and significance of '1' shows the eye closed state and significance of '0' shows the eye-open state. The first measured value is on the top i.e the data is arranged in descending order. Dataset consists of 14980 instances and 15 attributes with no missing values

C. Cognitive based dataset for BCI
The dataset was recorded was done on 3 people who have been inserted with grid electrodes for investigation of obstinate epilepsy. By doing so, an effort is made to obtain control over BCI implementing electrocorticographic (ECoG) signal of the left dorsolateral prefrontal cortex (DLPFC) [4]. Several runs of task were achieved [5] by considering power of high frequency in the range of 55Hz - 95Hz and FLm4 is the electrode used. Further 2 datasets of separate 4 min run are considered and is taken as a feedback task to proceed. Then patient is directed for counting the digits in reverse considering the from the digit visible from the monitor, thus making the pointer to be changed at the top end, relax, further move towards the bottom end. Task on Dataset 1 is completed openly, and Task on Dataset 2 is completed secretly

D. Prediction of finger flexion dataset
This data set contains brain signals from three subjects with epileptic patients, as well as the time courses to predict the flexion of each of five fingers. The experiment was done on every individual for 10 minutes [16]. Random clue was given on computer monitor for every 2 seconds signifying the finger to move. In the later stage numerous times the essential finger was moved by the subject. After every trial 2 seconds of rest period was taken to continue the next iteration [17].

III. METHODOLOGY

Fig. 1 shows the methodology used to achieve the objectives as mentioned. The datasets are primarily subjected to preprocessing in which cleaning of data is performed [11]. To obtain efficient results it is very much necessary to have a pure data with minimum noise or irrelevant data in the dataset. Preprocessing of dataset is performed to handle missing values in the data and the categorical variables in the dataset and to normalize the data if necessary [12].

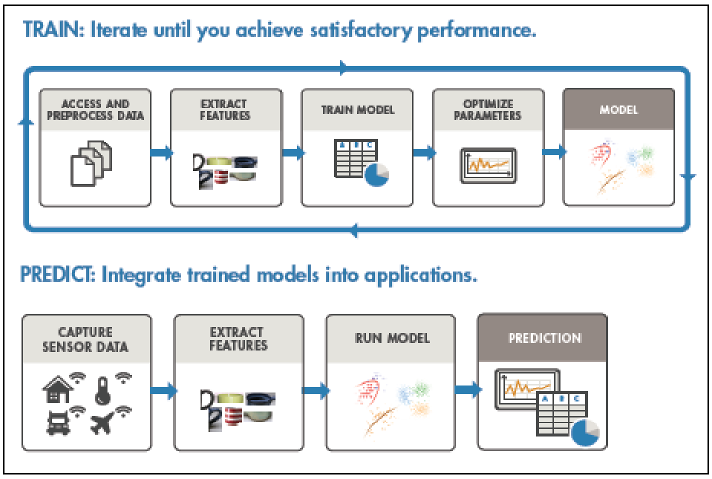

Fig. 1 Methodology

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 12, December 2021

DOI: 10.17148/IJARCCE.2021.101252

After preprocessing the data in the dataset [13], suitable procedures are implemented to the dataset to extract the best feature. Dataset dimension is condensed by feature extraction, which helps in decreasing the computational power of the learning process of the model and in retaining the most dependent data in the dataset [4]. Since the dataset dimension is reduced there will always be tradeoff between the efficiency and dimension reduction techniques, but it is always important to concentrate on computational power for the learning and prediction, Therefore the feature extraction technique places a key role in data modelling or training of the model [15]. Techniques like PCA, ICA, Wavelets and statistical approach are used for dimensionality reduction of the dataset.

A. Independent component analysis (ICA)

Blind source separation (BSS) problem can be overcome by considering an appropriate procedure called as Independent Component Analysis (ICA) [15]. Statistics adapted here is to first separate the signals into individual factors and consider as numerous resources. Therefore, extricated of multiple individual factors of the data is done involving this procedure. The number of individual factors in time series considered for multichannel information is m. Various diverse ICA algorithms are existing, of which three algorithms of ICA are predominantly accepted - FastICA, Infomax and JADE. Even though various several algorithms exist based on ICA, entire set is grounded on the thought that it consists of an equation given by x=As, where x can be a arbitrary vector where the contents contain mixture and ‘s’ is any arbitrary vector of elements and finally we have elements within a matrix indicated as ‘A’. Algorithms like Infomax, FastICA and JADE information is performed and studied. Considering EEG as analogy, the multiple sources here can be the multiple electrodes used to collect the various signals of brain activities at different positions of the scalp. So, investigation engaging ICA is meant for extracting the original signals from their mixtures. Additionally, the individual factors derived within the combinations of resources is disintegrated. The data derived through this procedure can be believed to be the output of brain actions and possibly is that it can be an artifact as well generated by Stimulation using Transcranial Magnetic Stimulation.

B. Principal component analysis (PCA)

Feature extraction of EEG signal can also be done by an alternative procedure Principal Component Analysis (PCA) which is generally employed. Karl Pearson initially in the year 1900 gave the concept of PCA and was first announced[15]. A plain, nonparametric procedure to extract evidence from disordered information and convey it in their cohesion and disparities is PCA. In PCA by turning round the data shifting is done and hence the focus on the data is found with utmost deviation. The initial element rotated to the direction of the data where the inconsistency is found to be the most is related to the initial axis. Therefore, the subsequent component will be vertical to the initial component with the utmost inconsistency and so on [18]. Apart from being capable of excellently extracting the feature, it as well supports the decline in the dimensionality and as result can lessen complicated information towards a smaller dimension and reveal concealed evidence inserted inside the information. So, a graphical illustration in examining and discovering relationships in dataset is done by PCA. Steps involved in working towards PCA algorithm is given below:

1. One elements mean is considered, and a covariance matrix C is prepared to resolve components of covariance and hence determine the Eigen value using $C - \lambda I$
2. The Eigen vectors are determined by employing eigen values and placed in matrix to examine the eigen values which is the largest.
3. The eigen vector found by choosing the biggest eigen value is the principal component.

C. Wavelet Transform

Waveform having oscillation of limited amplitude in a limited duration having zero as average value is defined as wavelet.

$$\varphi_{m,n}(t) = 1 \Big/ \sqrt{m\varphi\left(\frac{t-n}{m}\right)} \qquad (1)$$

The dilation parameter or scaling parameter is denoted as ‘m’ and translation parameter or position parameter is denoted as ‘n’ in (1).

The EEG signals are decomposed to obtain the coefficients at different levels using the discrete wavelet transform [10]. The wavelet decomposition of a signal based on multi resolution theory can be done using FIR filters.

D. SVM (Support Vector Machine)

The problem of classification and regression together can be solved by Support Vector Machines (SVM). The decision boundary aids in explaining the hyperplane. The decision plane differentiates the collection of objects belonging to various classes. If the objects are not distinguishable linearly occasionally, then complex mathematical functions named kernels help in differentiating the objects which are representatives of distinct classes.

The finest line or decision boundary is shaped to separate out to classes considering n-dimensional space and therefore the datapoint will effortlessly be classified newly to a category aptly in course of time. Thus, the right best decision border is the hyperplane.

IJARCCE

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 12, December 2021

DOI: 10.17148/IJARCCE.2021.101252

E. K-NN (K-Nearest Neighbour)

One of the algorithms considered for classification & regression is K Nearest Neighbour (KNN) Algorithm which utilizes a non-parametric method for classification because it does not consider any original data. Data points from the database are grouped into several classes based on same feature existing in each of the data points. Classification is done by assigning a number to K value.

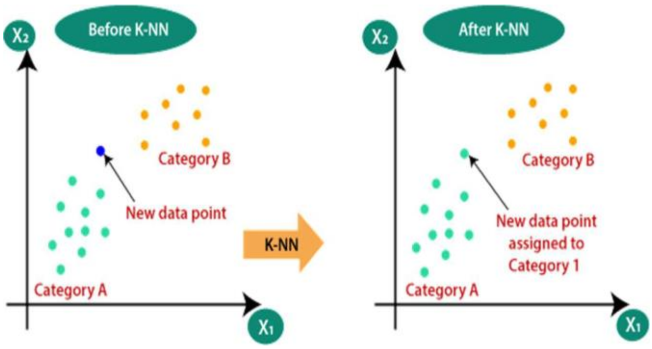

Fig. 2 K-NN

F. MLP(Multi-Layer Perceptron)

A class called as multilayer perceptron (MLP) is in feed forward Artificial neural network (ANN). Many a time it can be sometimes referred to as any feedforward ANN which occasionally sternly denote to associations created by multiple levels of perceptron’s (with threshold activation). Multilayer perceptron is every so very commonly mentioned by the name "vanilla" neural networks, particularly once they possess only one concealed layer.

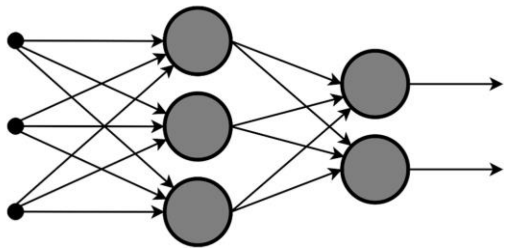

Fig. 3 Multi-Layer Perceptron

An input layer, a hidden layer and an output layer are nevertheless considered as 3 levels of points in MLP: Every single point is a neuron, exempting only the points at input where a nonlinear triggering work is employed. Backpropagation for training is applied by MLP which is a regulated understanding technique. MLP is recognized from a straight perceptron through multiple levels and not by linear stimulation and is also efficient of differentiating data that is not linearly separable.

G. Implementation on datasets

Let us consider the dataset Epilepsy Seizure detection using EEG. In this the feature extraction techniques such as statistical, PCA, wavelets for the dimensionality reduction of the dataset are applied, out of which achieved a highest efficiency in statistical technique and apply classifier such as Logistic Regression, SVM, KNN, random forest, Decision tree for classification of different classes, highest efficiency in both logistic regression and decision tree were achieved. This infers that the epileptic signals are more differentiable by statistical features.

The second dataset is Eye state detection. First the feature extraction techniques such as statistical, PCA, wavelets for the dimensionality reduction were applied but the achieved efficiency was not satisfactory, so dimension of the dataset was not reduced which yielded more efficiency when compared with the feature extraction technique and classifier such as Logistic Regression, SVM, KNN, Random Forest, Decision tree was applied, out of which KNN outperformed with highest efficiency. This shows that by reducing the dimension of the dataset efficiency drops which means there is more data dependency and hence dimension reduction should be done wisely.

The third dataset is Brain computer interfacing based on cognitive control, in this feature extraction technique such as PCA with different number of components considered such as 10 channels, 15 channels, 25 channels were considered for the dimensionality reduction [14]. Classifier such as Logistic Regression, KNN, random forest, Decision tree, MLP were implemented and MLP produced the highest efficiency. Infer that the data correlation between the channels varies by varying the components in PCA.

ISSN (O) 2278-1021, ISSN (P) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 10, Issue 12, December 2021
DOI: 10.17148/IJARCCE.2021.101252

The fourth dataset is prediction of finger movement. This dataset provides the clear picture about the correlation between the fingers during their flexion. The correlation was estimated by identifying the N best channel frequency pairs and applying regression to the selected channels.

IV. RESULTS AND CONCLUSION

A. Epilepsy Detection Result

Below shown are the results of the epileptic and non-epileptic signal classification for different techniques.

Table 1 Epilepsy results

Classification Algorithms	Statistical	PCA	Wavelets
Logistic Regression	100%	79.2%	80.43%
SVM	99.73%	93.88%	97.13%
K-NN	92.08%	93.33%	92.34%
Random Forest	99.97%	93.07%	96.52%
Decision Tree	100%	91.04%	94.63%

Hence, the result table can conclude that the statistic feature extraction is more suitable technique for the epilepsy detection since there is very minimal amount of data loss and data dependency is not lost when the statistical features are extracted. Whereas in other techniques even though there is dimension reduction of data, the data loss is not so minimal, and this is the reason for less efficient classification of signals for these extraction techniques.

B. Eye State Detection Results

Table 2 Eye state detection results

Classification Algorithms	Statistical	PCA	Wavelets	No Algorithm
Logistic Regression	53.82%	53.84%	53.84%	63.50%
SVM	53.82%	56.06%	56.49%	53.87%
K-NN	51.73%	56.72%	56.45%	95.79%
Random Forest	51.53%	56.09%	57.61%	88.94%
Decision Tree	51.60%	54.96%	55.36%	84.00%

From the result table it can concluded that the eye state detection is more accurate when the whole data is passed to KNN classifier, and no other feature extraction techniques were useful in increasing the efficiency of eye state detection since the data dependency is very high, reduction in dimensionality leads to higher data loss in this case.

C. Cognitive Dataset Results

Intentional objectives and correlated differences in brain activity is converted to actions and translated by Brain-computer interfaces (BCIs), thus patients suffering from acute paralysis is recommended an alternative process of interaction with their environment and monitor over it [8]. The person was given a task of counting a number backward which was displayed on screen and hence by doing so, the cursor was sent to upper target and made the individual person in resting positioning, henceforth the cursor moves to the bottom level.

Table 3 Cognitive dataset results

Classification algorithms	Pre-processing Techniques		Dimension Reduction Technique			No Pre-processing
Logistic Regression	64%	60%	PCA			66%
K-NN	97%	98%	10n	15n	25n	98%
Random Forest	87%	87%	57%	59%	59%	66%
Decision Tree	79%	79%	84%	89%	93%	87%
MLP	99%	99%	98%	98%	98%	99%

D. Prediction on Finger Flexion Results

The objective of this experimentation is to discover correlation between finger moments and do a detailed examination of data, to find significant associations inside data, and construct a prediction model.

IJARCCE

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 12, December 2021

DOI: 10.17148/IJARCCE.2021.101252

Table 4 Finger flexion correlation

Finger	Correlation
Finger-1	0.625
Finger-2	0.620
Finger-3	-0.009
Finger-5	0.29
Average (1,2,3,5)	0.38

E. Snapshots of webpage

A webpage as in Fig. 4 was developed for the prediction of the classes for different combination of datasets, feature extractions and classifiers. HTML and CSS were used for creating the forms for selection of combinations by the users.

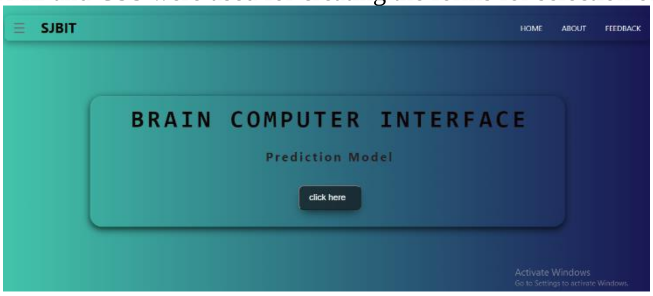

Fig. 4 Landing page of website

The interface takes the user input in the form of data and passes it to the model to do the further predictions as in Fig. 5. The trained ML model are connected to the interface by means of flask web framework, which helps in collecting the HTML form data and processing of the prediction by the trained model

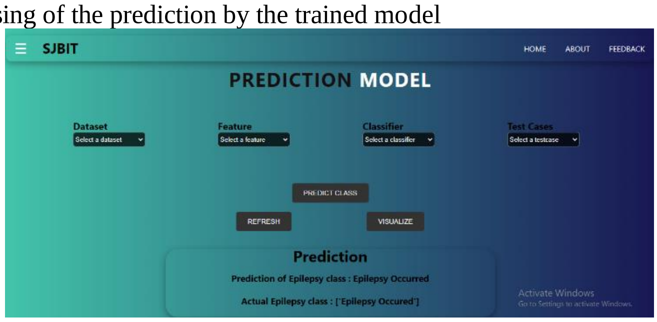

Fig. 5 Prediction model webpage

To display the prediction on to the web page, jinja code was used. Chart.js was used to visualize the input test signals as in Fig. 6

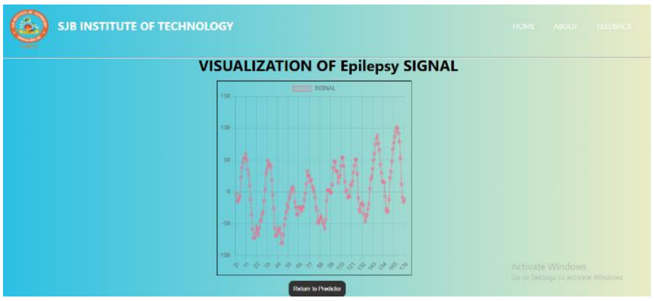

Fig. 6 Visualization of signals in webpage

The analysis of the signals in eye state and epilepsy datasets conclude that two classes can be classified quite accurately, but when it comes to multiclass datasets, few extractions and selection of feature techniques should be used to wisely reduce the data dimension and increase the data dependency. In several lifesaving purposes, human eye state recognition will be able to widely be adopted, except for incidents like vehicle collision, lack of sleep, an individual positioned at critical safety areas and so on., where it is typically booked a case owing to dearth of focus or lack of concentration

The predictions of the model will be implemented as commands to monitor the machine and for other BCI purposes by adding a real time brain signal input feature to the user interface. Pprominence on the assessment of any phases are vital

ISSN (O) 2278-1021, ISSN (P) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 10, Issue 12, December 2021
DOI: 10.17148/IJARCCE.2021.101252

to make the system work, along with uncomplicated evaluation of its triumph is the improvement and user-centrered estimation of engineered BCI applications.

Applications of BCI technology to identify individual human perception, impact, action, and several qualities of reasoning and behaviour.

REFERENCES

[1] Rabie A. Ramadan and Athanasios V. Vasilakos."Brain Computer Interface:Control Signals Review". Neurocomputing, Volume 223, 5 Feb 2017, pp 26–44.

[2] Levin Kuhlmann, Philippa Karoly and Dean R. Freestone. "Epilepsyecosystem.org: crowdsourcing reproducible seizure prediction with long-term human intracranial EEG". Brain 2018: 141,pp 2619–2630.

[3] Amirali Vahid, Moritz Muckschel, Sebastian Stober."Applying deep learning to single-trial EEG data provides evidence for complementary theories on action control". Communication Biology.

[4] Guruprasad Madhale Jadav, Jonatan Lerga and Ivan Stajduhar."Adaptive filtering and analysis of EEG signals in the time-frequency domain based on the local entropy".EURASIP Journal on Advances in Signal Processing

[5] Raju Vishwakarma, Hazim Khwaj, Varad Samant and Prajyot Gaude."EEG Signals Analysis and Classification for BCI Systems: A Review". 2020 International Conference on Emerging Trends in Information Technology and Engineering.

[6] Rabie A. Ramadan, S. Refat, Marwa A. Elshahed and Rasha."Basics of Brain Computer Interface". Springer International Publishing Switzerland 2015.

[7] Swati Vaid, Preeti Singh and Chamandeep Kaur."EEG Signal Analysis for BCI Interface: A Review". 2015 Fifth International Conference on Advanced Computing & Communication Technologies.

[8] Georg E. Fabiani, Dennis J. McFarland, Jonathan R. Wolpaw, and Gert Pfurtscheller."Conversion of EEG Activity into Cursor Movement by a Brain–Computer Interface (BCI)".IEEE Transactions on Neural Systems and Rehabilitation Engineering, Vol. 12, no. 3, Sep 2004.

[9] J Satheesh Kumar and P Bhuvaneswari."Analysis and Electroencephalography (EEG) Signals and Its Categorization-A Study". International Conference on Modeling, Optimization and Computing (ICMOC 2012).

[10] Er. Jasjeet Kaur and Er. Amanpreet Kaur."A Review on Analysis of EEG Signals". 2015 International Conference on Advances in Computer Engineering and Applications (ICACEA) IMS Engineering College, Ghaziabad, India.

[11] Eduardo Lopez-Larraza, Thiago C. Figueiredoa, Ainhoa Insausti-Delgado."Event-related desynchronization during movement attempt and execution in severely paralyzed stroke patients: An artifact removal relevance analysis". NeuroImage: Clinical 20 (2018) 972–986.

[12] Andrzej Majkowski, Marcin Kolodziej and Remigusz J. Rak."Implemantation of Selected EEG Signal Processing Algorithms in Asynchronous BCI". Institute of theory of Electrical Engineering, Measurement of Information Systems Warsaw University of Technology Poland.

[13] Wan Amirah W Azlan and Yin Fen Low."Feature Extraction of Electroencephalogram (EEG) Signal – A Review". 2014 IEEE Conference on Biomedical Engineering and Sciences, 8 - 10 December 2014, Miri, Sarawak, Malaysia.

[14] Antoine Gaume, Gerard Dreyfus and François-Benoit Vialatte."A cognitive brain–computer interface monitoring sustained attentional variations during a continuous task". Cognitive Neurodynamics (2019) 13:257–269.

[15] Susmita Ray."A Quick Review of Machine Learning Algorithm". 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (Com-IT-Con), India, 14th -16th Feb 2019.

[16] Hohyun Cho, Minkyu Ahn, Sangtae Ahn, Moonyoung Kwon and Sung Chan Jun."EEG datasets for motor imagery brain–computer interface". Giga Science, 6, 2017, 1–8.

[17] Prashant Lahane, Jay Jagtap, Aditya Inamdar, Nihal Karne and Ritwik Dev."A review of recent trends in EEG based Brain-Computer Interface". Second International Conference on Computational Intelligence in Data Science (ICCIDS-2019).

[18] K. Mahantesh and R. Chetana, “Detection of epileptic seizures in EEG- inspired by machine learning techniques,” Intelligent Computing and Communication, Advances in Intelligent Systems and Computing, vol. 1034, pp. 443–450, 2019.