International Journal of Trend in Scientific Research and Development (IJTSRD)
International Open Access Journal
ISSN No: 2456 - 6470 | www.ijtsrd.com | Volume - 2 | Issue – 1

A review on Machine Learning Techniques for Neurological disorders estimation by Analyzing EEG Waves

Vijaykumar Janga
Reaserch Scholor, Acharya Nagarjuna Univerity,
Nagarjuna Nagar, Guntur

Dr Eedara Srinivasareddy
Principal, ANU College of Engg & Technology,
Acharya Nagarjuna University, Guntur

ABSTRACT

With the fast improvement of neuroimaging data acquisition strategies, there has been a significant growth in learning neurological disorders among data mining and machine learning communities. Neurological disorders are the ones that impact the central nervous system (including the human brain) and also include over 600 disorders ranging from brain aneurysm to epilepsy. Every year, based on World Health Organization (WHO), neurological disorders affect much more than one billion people worldwide and count for up to seven million deaths. Hence, useful investigation of neurological disorders is actually of great value. The vast majority of datasets useful for diagnosis of neurological disorders like electroencephalogram (EEG) are actually complicated and poses challenges that are many for data mining and machine learning algorithms due to their increased dimensionality, non stationarity, and non linearity. Hence, an better feature representation is actually key to an effective suite of data mining and machine learning algorithms in the examination of neurological disorders. With this exploration, we use a well defined EEG dataset to train as well as test out models. A preprocessing stage is actually used to extend, arrange and manipulate the framework of free data sets to the needs of ours for better training and tests results. Several techniques are used by us to enhance system accuracy. This particular paper concentrates on dealing with above pointed out difficulties and appropriately analyzes different EEG signals that would in turn help us to boost the procedure of feature extraction and enhance the accuracy in classification. Along with acknowledging above issues, this particular paper proposes a framework that would be useful in determining man stress level and also as a result, differentiate a stressed or normal person/subject.

Keywords: Electroencephalogram (EEG), Emotion recognition, Stress, Machine learning techniques

I. INTRODUCTION

Mental disorders or neurological disorders are increasing at high pace in the world. As per WHO, one among four people in the world will be affected by mental or neurological disorders at some point of time in their life. Neurological disorders are going to be second leading cause of global disease burden by year 2020, lagging behind ischemic heart illness but leading all the other diseases [1]. The increase in the number of professionals who treat the mental illness is very less as compared to the growth in number of people who are suffering from mental problems. Mental health diagnoses involve steps like specially designed interviews about symptoms and medical data and sometimes physical examination of the patient. Several psychological tests may also be conducted to make sure the symptoms are due of mental health problems and not because of any other disease. Similarity in the symptoms of several mental health disorders has made diagnosis complicated task. Diagnoses of mental health problems in children are far more difficult than diagnosing them in adults. Therefore one needs to be careful to diagnose the mental health disorders with accuracy. It's known that psychiatric/neurological disorders affect brain function and structure. However, to date the translation of neuroimaging research findings into diagnostic tools have been very limited due to lack of adequate analysis tools. In the last years there has been a substantial increase in the use of machine learning/pattern recognition approaches to analyze neuroimaging data.

Artificial Intelligence can enable the computer to think. Computer is made much more intelligent by AI. Machine learning is the subfield of AI study. Various researchers think that without learning, intelligence cannot be developed. There are many types of Machine Learning Techniques that are shown in Figure 1. Supervised, Unsupervised, Semi Supervised, Reinforcement, Evolutionary Learning and Deep Learning are the types of machine learning techniques. These techniques are used to classify the data set.

1) Supervised learning: Offered a training set of examples with suitable targets and on the basis of this training set, algorithms respond correctly to all feasible inputs. Learning from exemplars is another name of Supervised Learning. Classification and regression are the types of Supervised Learning.

Classification: It gives the prediction of Yes or No, for example, "Is this tumor cancerous?", "Does this cookie meet our quality standards?"

Regression: It gives the answer of "How much" and "How many".

2) Unsupervised learning: Correct responses or targets are not provided. Unsupervised learning technique tries to find out the similarities between the input data and based on these similarities, un-supervised learning technique classify the data. This is also known as density estimation. Unsupervised learning contains clustering [1]. Clustering: it makes clusters on the basis of similarity.

3) Semi supervised learning: Semi supervised learning technique is a class of supervised learning techniques. This learning also used unlabeled data for training purpose (generally a minimum amount of labeled-data with a huge amount of unlabeled-data). Semi-supervised learning lies between unsupervised-learning (unlabeled-data) and supervised learning (labeled-data).

4) Reinforcement learning: This learning is encouraged by behaviorist psychology. Algorithm is informed when the answer is wrong, but does not inform that how to correct it. It has to explore and test various possibilities until it finds the right answer. It is also known as learning with a critic. It does not recommend improvements. Reinforcement learning is different from supervised learning in the sense that accurate input and output sets are not offered, or suboptimal actions clearly précised. Moreover, it focuses on on-line performance.

5) Evolutionary Learning: This biological evolution learning can be considered as a learning process: biological organisms are adapted to make progress in their survival rates and chance of having off springs. By using the idea of fitness, to check how accurate the solution is, we can use this model in a computer [2].

6) Deep learning: This branch of machine learning is based on set of algorithms. In data, these learning algorithms model high-level abstraction. It uses deep graph with various processing layer, made up of many linear and nonlinear transformation.

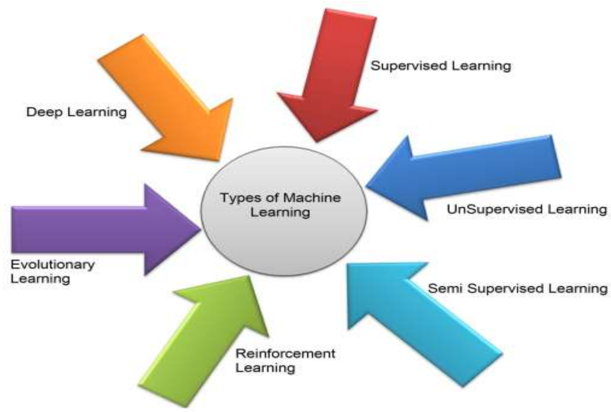

Figure 1: Types of machine learning techniques

Electroencephalography (EEG) is a monitoring method which can help to record the electrical activity of the brain. This electrical activity can lead us to better understand the human brain and how it functioning. Brain Computer Interface combine hardware and software communication system that permits cerebral activity alone to control computers and other devices.

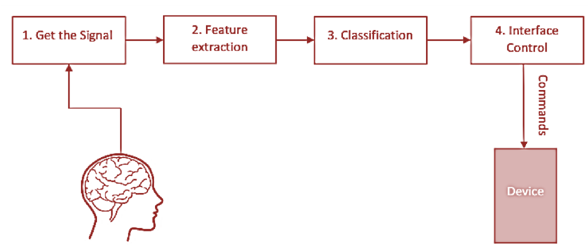

Figure 2: Design of a BCI system

BCI enables to interact with the surroundings, without the involvement of peripheral nerves and muscles, by using control signals generated from electroencephalographic activity. There are several stages to do so as depicted in Figure.2 and its process shown below:

1) **Get the Signals**: Capture the brain signals and make noise reduction and preprocessing the signals in order to be able to process it in more a convenient

way

2) **Feature Extraction:** Identifies discriminative information in the brain signals that have been recorded. This can be a challenging job, because of the many mixed signals with large number of sets activity in the brain that overlap in time and space, and we don't want to loss information
3) **Classification**: classify the signals to achieve pattern recognition in order to decipher the user's intentions
4) **Interface Control**: Translate the classified signals into the user desired commands for any kind of device such as a computer.

II. RELATED WORKS

Masri RY and Jani HM [5] offered the mental health Diagnostic Expert System for the assistance of psychologists to diagnose and treat their mental patients. Three artificial techniques viz., Fuzzy Logic, Rule-Based Reasoning and Fuzzy Genetic Algorithm were applied in diagnosing and suggesting the treatment plans. Luxton et al. [7] analyzed the use of artificial intelligence for psychological task.

Razzouk D et al. [8] developed the decision supporting system for diagnosis of schizophrenia having accuracy up to 66-82%. Chattopadhyay S et al. [9] developed a neuro-fuzzy approach for categorizing of adult depression. The supervised Adaptive Network Based Fuzzy Inference System and Back Propagation Neural Network and unsupervised Self Organizing Map neural network learning techniques were utilized and compared. It was observed that Adaptive Network Based Fuzzy Inference System, a hybrid system performed far better than Back Propagation Neural Network.

Basavappa SR et al. [10] applied depth first search algorithm with the backward search approach for diagnosing dementia. An expert system was developed by them taking in consideration patient's behavior, cognition, emotions and the results of neuropsychological tests. Rahman et al. [11] compared several classification techniques; Multilayer Perceptron, Bayesian Network, Single Conjunctive Rule Learning, Decision Trees, Neuro-Fuzzy Inference System and Fuzzy Inference Systems using various data mining softwares like TANAGRA ,WEKA and MATLAB for diagnosing diabetes. They observed that accuracy levels are different for different techniques on different accuracy measures such as Kappa Statistic and Error rates.

Gomuła, Jerzy et al. tried finding efficient techniques for the classification of MMPI profiles of patients having mental problem. They found that Attribute Extension methodology improves classification accuracy in case of discreatised data[12]. Anchana Khemphila,Veera Boonjing applied Multi-Layer Perceptron with Back Propagation Learning for diagnosing Parkinson's disease efficiently with selected attributes. Information Gain from all attributes is taken as a measure for the reduction of attributes [13]. Pirooznia Mehdi et al. [14] used data mining techniques to find Genome wide Association in Mood Disorders. Six classifiers Support Vector Machine, Bayesian Network, Logistic Regression, Radial-Basis Function, Random Forest and Polygenic Scoring method were being compared. It was found that a simple polygenic score classifier performed much better than others and they also found that all classifiers performed worse with small number of Single Nucleotide Polymorphisms in brain expressed set compared to whole genome set.

As it can be seen from the earlier sections, a wide range of research studies has been done for EEG artifacts removal. Methods that have been proposed can be divided into manual, semi-automatic and automatic. Manual and semi-automatic methods require expert observations to identify artifacts in EEG signal. On the other hand, automatic methods require predefined threshold value. In the past few years, machine learning techniques have been advanced significantly and used in pattern identification and classification problems. Table 1 presents a summary of the papers based on the different machine learning algorithms presented earlier in this paper. Table 1 shows that the SVM is the mostly used method and different approaches of SVM are applied to classify artifacts in EEG signal. Gaussian kernel and radial basis function (RBF) are found most appropriate approaches for EEG artifacts.

Table 1 Different machine learning algorithms

Machine Learning Technique	Associated Methods	References
Support Vector Machine	ICA, BSS, Autoregressive model	(Bartels et al., 2010; Chin-Teng et al., 2012; Gao, Yang, et al., 2010; Halder et al., 2007; Hsu et al., 2012; Lawhern et al., 2012; Lawhern et al., 2013a; O'Regan et al., 2013; O'Regan & Marnane, 2013; Phothisonothai et al., 2012; Shi Yun et al., 2009; Shi-Yun et al., 2008; Singla et al., 2011; Tangermann et al., 2009; Winkler et al., 2011; Wu et al., 2009)
Artificial Neural Network	ICA, Spectral analysis	(Chin-Teng et al., 2012; Jafarifarmand & Badamchizadeh, 2013; Junfeng et al., 2009; Marquez L & Munoz G, 2013; Nguyen et al., 2012; Singla et al., 2011; Sovierzoski et al., 2009)
Fuzzy Inference system	Differential Evolution Adaptive Noise Cancellation	(Kezi Selva Vijilal et al., 2007; Sheniha et al., 2013)
Clustering	Kurtosis	(Nicolaou & Nasuto, 2007; Patidar & Zouridakis, 2008; Yuan et al., 2012)
K-NN	Polynomial fitting, Hjort descriptor	(Aydemir et al., 2012; Gao, Lin, et al., 2010; Pourzare et al., 2012)
Bayesian Model	Spectral power	(Schetinin & Maple, 2007)
Genetic programming	Power spectral analysis, kurtosis	(Fairley et al., 2010; Poli et al., 2011)

III. SYSTEM EXEMPLARY

1. System description

In any classification system, feature selection and extraction is main and important phase toward successful classification system. In our case it's hard to think directly about which features and which classifiers to use in order to get the best results. The diversities are mainly in aspects of EEG artifacts, experiment environment, techniques of data preprocessing and feature selection. Due to all this factors, it is not easy to compare and chose the method which can be said as the best classifier. Hence, there is always room for the development of better classifier suitable for specific application.

OUR APPROACH:

Firstly the problem, the diagnosis of basic psychological health was identified followed by knowing the psychological health disorders that are often found in patients. A list of machine learning techniques for diagnosis of five most common psychological health disorders effectively if the symptoms of the patient are provided as input. The data sets of 25 attributes containing the class type labels that are found. The set includes these attributes: Age, Family, History, Pregnancy Complication, Delayed Speech, Under Medication, Academic Performance, Relationship Formation, Behavioural Problem, Concentration, Restless, Seizures, Learning Difficulty, Attention Aroused, Attention Sustained, CBCL Score, IQ Test Score, ADHD Positive, ODD Positive, Manic

Episode Test Score, Major Depressive Episode, General Anxiety Disorder, CDI Score, PDD Score, Autism Score and Problem Since only few attributes are relevant to classify and predict the problem, Best First Search technique is used to eliminate redundant and irrelevant attributes. This will also help in achieving more accuracy.

The performance analysis of the three classification algorithms has been carried out with common dataset applying WEKA tool or Matlab tool. The classifiers were executed by including selected attributes (13) only using feature selection method. WEKA tool bestows with the various measures for understanding the classification. Among the number of measures, the three measures which are very important for the comparison of the accuracy level of different classifiers are Kappa Statistics, ROC Area and Accuracy.

Tools used: In order to process the recorded signals, we need to use some softwares as a platform.

1) **EDF browser:** EDF Browser is a free open-source, multiplatform viewer and toolbox for time series storage files like EEG data. European Data

Format (EDF) is a standard file format designed for exchange and storage of medical time series. It offers a graphic visualization of the signal, as well as an integrated list of trigger marks present in the file. It also provides filtering functionalities, power on the frequency bands computation, as well as the possibility of down-sampling the signal. This program converts all the signals in an EDF to a plain ASCII text-file. Internally it includes a header and one or more data records. The data records contain consecutive fixed duration epochs of the poly-graphic recording. The header contains some general information (patient identification, start time…) and technical specs of each signal (calibration, sampling rate), coded as ASCII characters. A screenshot from EDF browser is shown in figure 3.

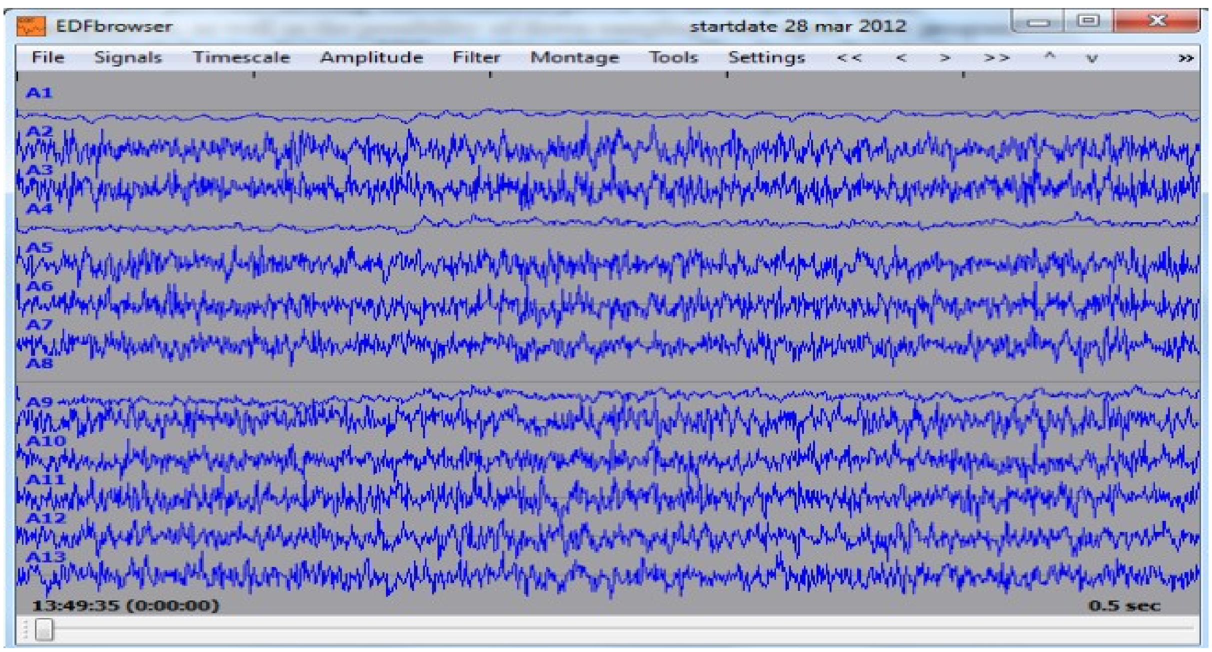

Figure 3: EDFbrowser [25]

MATLAB: MATLAB is a powerful tool, especially with the signal processing toolbox. It integrates computation, visualization, and programming environment. Furthermore, MATLAB is a modern programming language environment: it has sophisticated data structures, contains built-in editing and debugging tools, and supports object-oriented programming. MATLAB has functionality to analyze data, develop algorithms, and create models and applications.

The language tools and built-in math functions enable you to explore multiple approaches and reach a solution faster than with spreadsheets or traditional programming languages. These factors make MATLAB an excellent tool for teaching and research. It provides vast range of different functionalities for analyzing and processing EEG data filtering, time/frequency transforms, feature extraction etc. The Figure.4 is the screenshot from MATLAB.

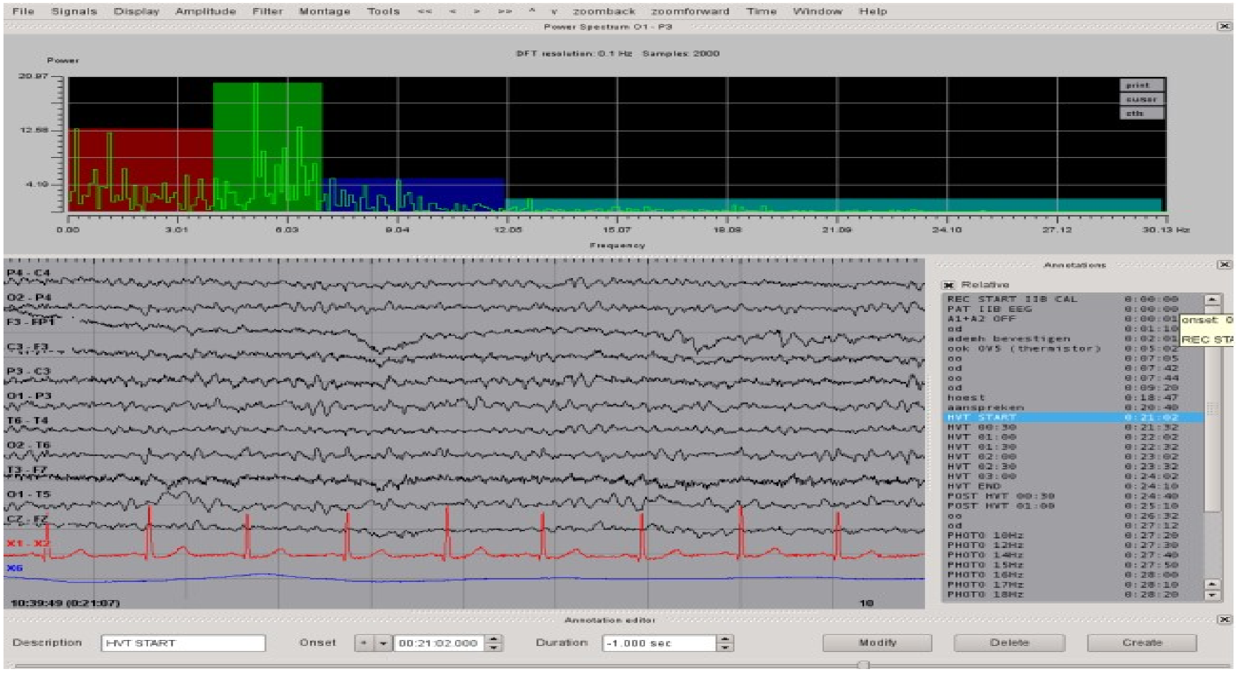

Figure.4 Manipulating EEG Signals and their Annotations [24]

WEKA Analysis: Waikato Environment for Knowledge Analysis (WEKA) is an open-source collection of machine learning algorithms for data mining tasks. The software is a widely accepted standard in the field and is commonly used in a variety of applications, ranging from biomedical to financial data analysis. WEKA is written in the Java programming language and is normally run under a Java Virtual Machine. Each machine learning algorithm implementation requires the data to be present in its own format, and has its own way of specifying parameters and output. We use Explorer window for our project as shown in Figure.5

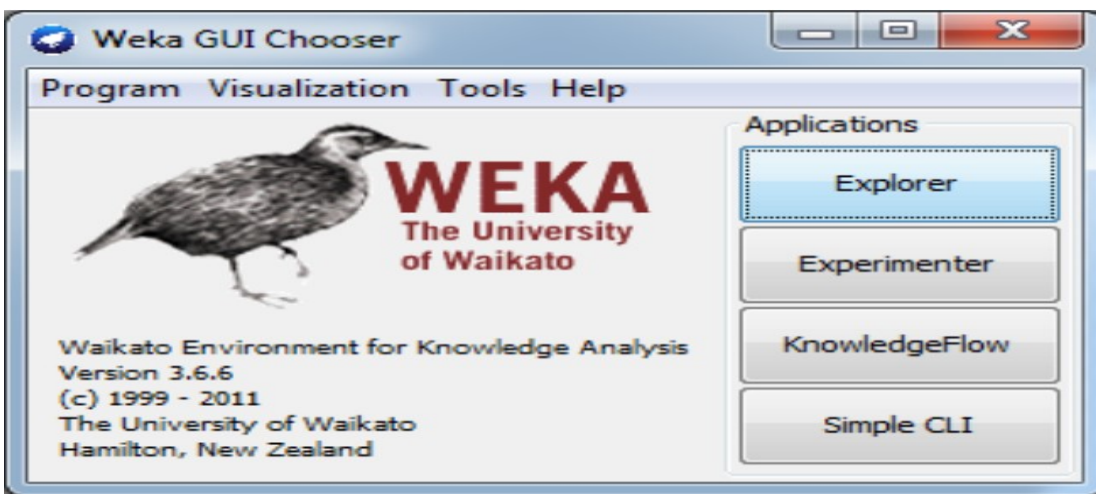

Figure.5 Waikato GUI Chooser [25]

A. Performance and robustness measures

The study of different types of oscillations and rhythmicities of the brain and their relation with different pathologies and functions keep the attention of researchers since the beginnings of EEG measuring. Brain oscillations were divided in frequency bands that have been related with different brain states, functions or pathologies. The characteristic oscillations are (Table II):

- Delta rhythms (0.5–3.5 Hz) are characteristic of deep sleep stages; delta oscillations with certain specific morphologies, localizations and rhythmicities are correlated with different pathologies,

- Theta rhythms (3.5–7.5 Hz) are enhanced during sleep and they play an important role in infancy and childhood; in the awake adult, high theta activity is considered abnormal and it is related with different brain disorders,
- Alpha rhythms (7.5–12.5 Hz) appear spontaneously in normal adults during wakefulness, under relaxation and mental inactivity conditions; they are best seenwith eyes closed and most pronounced in occipital locations,
- Beta rhythms (12.5–30 Hz) are best defined in central and frontal locations, they have less amplitude than
- alpha waves and they are enhanced upon expectancy states or tension, gamma rhythms (30–60 Hz) are generally not of major interest with regard to the surface EEG.

Table II: The characteristic brain oscillations.

Wave	frequency	voltage	condition
delta	0.5–3.5 Hz	10 mV	deep sleep
theta	3.5–7.5 Hz	adults: 10 µV kids: 50 µV	light sleep, drowsy
alpha	7.5–12.5 Hz	adults: 50 µV kids: 75 µV	relaxed
beta	12.5–30 Hz	10–20 µV	excited

CONCLUSION

In medical domain, numbers of expert systems are available to predict diseases at very early stage to make the treatment effective and efficient. In the similar manner, expert systems have been developed in psychological health sector for predicting the mental health problems at early stage. Since number of machine techniques are present for building expert systems, analysis of the techniques and their comparison for identifying the best technique which suits domain. This paper presents a literature review of machine learning algorithms that are frequently used psychological health sector handling. This article provides an overview of how certain machine leaning techniques have been applied in handling different EEG artifacts. From the study, it is revealed that a large number of automatic and semi-automatic methods are available for EEG artifacts removal. However, the usage of machine learning algorithms is limited. It is also found that machine learning algorithms provide better classification accuracy than other approaches. Moreover, comparison of different techniques is also studied and in several studies it is suggested that SVM is better classifier than other classification methods. Finally, the survey leaves us with focus on hybrid approaches i.e., using several machine learning algorithms.

REFERENCES

1. http://blogs.fortishealthcare.com/mental-health-india-wake-up-call/
2. Emotiv - Brain Computer Interface: http://www.emotiv.com
3. Bartels, G., Shi, L. C., & Lu, B. L. (2010). Automatic artifact removal from EEG - a mixed approach based on double blind source separation and support vector machine. Conference proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference, 2010, 5383-5386.
4. MY Rozita, JM Hajar, et al. Employing artificial intelligence techniques in Mental Health Diagnostic Expert System. In Computer & Information Science (ICCIS), International Conference 2012; 1: 495-449.
5. MR Sumathi, B Poorna, Prediction of Mental Health Problems among Children Using Machine Learning Techniques. International Journal of Advanced Computer Science and Applications 2016; 7.
6. LD David, Artificial intelligence in psychological practice: Current and future applications and implications. Professional Psychology: Research and Practice 2014; 45: 332-339.

7. D Razzouk, JJ Mari, et al. Decision support system for the diagnosis of schizophrenia disorders. Brazilian Journal of Medical and Biological Research 2006; 39: 119-128.

8. C Subhagata, K Preetisha, et al. An automated system to diagnose the severity of adult depression. In Proceedings of Second International Conference on Emerging Applications of Information Technology, IEEE Computer Society and Conference Publishing Services, Kolkata, India 2011; 121–124.

9. SR Basavappa, SL Rao, et al. Expert system for dementia or depression diagnosis. NIMHANS Journal 1996; 14: 99-106.

10. RM Rashedur, A Farhana, Comparison of various classification techniques using different data mining tools for diabetes diagnosis. Journal of Software Engineering and Applications 2013; 6: 85-97.

11. G Jerzy, P Krzysztof, et al. Classification of MMPI profiles of patients with mental disorders–experiments with attribute reduction and extension. Rough Set and Knowledge Technology 2010; 411-418.

12. K Anchana, B Veera, Parkinsons disease classification using neural network and feature selection. World Academy of Science, Engineering and Technology 2012; 64: 15-18.

13. P Mehdi, F Seifuddin, et al. Data mining approaches for genome-wide association of mood disorders. Psychiatric genetics 2012; 22: 55-61.

14. Chin-Teng, Lin, Yu-Kai, Wang, & Shi-An, Chen. (2012, 10-15 June 2012). A hierarchal classifier for identifying independent components. Paper presented at the Neural Networks (IJCNN), The 2012 International Joint Conference on.

15. Gao, JunFeng, Yang, Yong, Lin, Pan, Wang, Pei, & Zheng, ChongXun. (2010). Automatic Removal of Eye-Movement and Blink Artifacts from EEG Signals. Brain Topography, 23(1), 105-114. doi: 10.1007/s10548-009-0131-4.

16. Halder, Sebastian, Bensch, Michael, Mellinger, Jurgen, Bogdan, Martin, Kubler, Andrea, Birbaumer, Niels, & Rosenstiel, Wolfgang. (2007). Online Artifact Removal for Brain-Computer Interfaces Using Support Vector Machines and Blind Source Separation. Computational Intelligence and Neuroscience, 2007.

17. Hsu, Wei-Yen, Lin, Chao-Hung, Hsu, Hsien-Jen, Chen, Po-Hsun, & Chen, I. Ru. (2012). Wavelet-based envelope features with automatic EOG artifact removal: Application to single-trial EEG data. Expert Systems with Applications, 39(3), 2743-2749.

18. Lawhern, Vernon, Hairston, W. David, McDowell, Kaleb, Westerfield, Marissa, & Robbins, Kay. (2012). Detection and classification of subject-generated artifacts in EEG signals using autoregressive models. Journal of Neuroscience Methods, 208(2), 181-189.

19. Lawhern, Vernon, Hairston, W. David, & Robbins, Kay. (2013a). DETECT: A MATLAB Toolbox for Event Detection and Identification in Time Series, with Applications to Artifact Detection in EEG Signals. PLoS ONE, 8(4), e62944. doi: 10.1371/journal.pone.0062944.

20. O'Regan, Simon, Faul, Stephen, & Marnane, William. (2013). Automatic detection of EEG artefacts arising from head movements using EEG and gyroscope signals. Medical Engineering & Physics, 35(7), 867-874.

21. O'Regan, Simon, & Marnane, William. (2013). Multimodal detection of head-movement artefacts in EEG. Journal of Neuroscience Methods, 218(1), 110-120.

22. Phothisonothai, M., Fang, Duan, Tsubomi, H., Kondo, A., Aihara, K., Yoshimura, Y., Watanabe, K. (2012, 5-7 Dec. 2012). Artifactual component classification from MEG data using support vector machine. Paper presented at the Biomedical Engineering International Conference (BMEiCON), 2012.

23. Ahmad Tauseef Sohaib & Shahnawaz Qureshi, Empirical study of various machine learning technique for classifying emotional state from EEG Data, Blekinge Institute of Technology, Sweden, 2012.

24. http://www.teuniz.net/edfbrowser/EDFbrowser%20manual.html [Accessed: 13th May 2013].