

# **BRAIN COMPUTER INTERFACES EMPLOYING MACHINE LEARNING METHODS : A SYSTEMATIC REVIEW**

Sejal Vyas

Department of Computer Science, Stevens Institute of  
Technology, Hoboken, NJ, USA

## **ABSTRACT**

*Research under the field of Brain Computer Interfaces is adapting various Machine Learning and Deep Learning techniques in recent times. With the advent of modern BCI, the data generated by various devices is now capable of detecting brain signals more accurately. This paper gives an overview of all the steps involved in the process of applying Machine Learning as well as Deep Learning methods from Data Acquisition to application of algorithms. It aims to study techniques currently employed to extract data, features from brain data, different algorithms employed to draw insights from the extracted features, and how it can be used in various BCI applications. By this study, I aim to put forward current Machine Learning and Deep Learning Trends in the field of BCI.*

## **KEYWORDS**

*Brain Computer Interface (BCI), Machine Learning, Deep Learning*

## **1. INTRODUCTION**

Brain Computer Interface (BCI) is a system which is used to perform certain tasks based on the processing of brain activities. These systems act as a medium which facilitates human interaction with external devices. To give a very high-level perspective, data (for example brain waves) is recorded. This brain data is fed to BCI which in turn generates control commands to operate applications. Thus, BCI systems are used to control say for example a robotic arm by a quadriplegic subject by means of her brain waves. While BCI applications are mostly used for assistive technologies, they project a range of applications [1] in Gaming and Entertainment, Training and Education, Cognitive Improvement, User-State Monitoring, Safety and Security, Device Control, and Evaluation. Machine learning comes into picture when brain data is analysed. Over the course of this paper, I would also study some applications of Deep Learning techniques in BCI. The goal of this paper is to shed light on currently employed machine learning and deep learning methods in this field. This paper first describes the various kinds of data involved in such systems. Depending on the type of data, pre-processing and feature extraction process is described. In addition to that, a survey is presented on various Machine Learning and Deep Learning trends in this field. This study will help in getting a clear idea of which techniques are currently being employed in various BCI applications. In conclusion, I put forward a brief comparison which will help in selecting appropriate machine learning and deep learning techniques for different use cases.

## 2. DATA ACQUISITION

To obtain brain data, various technologies are employed which can be invasive or non-invasive. These technologies as listed in [1] are, Electroencephalography (EEG), Intracortical Electrodes (ICE), functional Near-Infrared Spectroscopy (fNIRS), functional Magnetic Resonance Imaging (fMRI), Magnetoencephalography (MEG), and Electroencephalography (EEG).

Electroencephalography (EEG), is a method where electrode grid is placed on the surface of the brain to record electrical activity from the cortex. Intracortical Electrodes (ICE) are devices that are inserted inside the brain, deep into the grey matter to record brain signals. Since both of these methods are invasive, the strength of the signals recorded is strong. The type of data for both these techniques is functional data (usually a function of time). Functional Near-Infrared Spectroscopy (fNIRS), which is a non-invasive technique, measures the signals based on the hemodynamic responses associated with the neuron behaviour. Functional Magnetic Resonance Imaging (fMRI) is also a non-invasive technique which measures brain activities by detecting changes associated with blood flow. fNIRS and fMRI both employ neuroimaging techniques for extracting brain activities. Image data is obtained by both these techniques. Magnetoencephalography (MEG) is a non-invasive technique which measures magnetic fields generated by brain. The data obtained by this technique is functional data. Electroencephalography (EEG), is a non-invasive technique, which measures voltage fluctuations resulting from ionic movements of neurons. The data obtained from this is also functional data. Out of all the techniques, EEG is the most widely used technique [1]. The data obtained from all the technologies is either functional data or image data.

Data thus recorded from various data extraction techniques is further processed and features are obtained from it. For cases where data acquisition directly from subjects is not possible, open source datasets are used and features extraction techniques are applied on them.

## 3. PRE-PROCESSING AND FEATURE EXTRACTION

Pre-processing is required if the data obtained is of functional type or of image type. Let us first consider functional data. The signals are measured against time thus this data may also be called time series data. Recorded signals are contaminated at the acquisition stage. Such signals are called artifacts [2]. There are two types of artifacts as mentioned in [2], physiologic and extra-physiologic. Extra-physiologic artifacts are caused due to external factors like devices or environment while recording signals. Physiologic artifacts are caused due to human body other than brain [2]. As classified in [2], extra-physiologic artifacts involve alternating current artifacts, electrode artifacts, artifacts caused due to movements in recording environment, and artifacts produced due to interference with other devices. Physiologic artifacts include signals generated by electrooculography, electrocardiogram, pulse, skin, tongue movement (also known as glossokinetic artifact), and electromyogram. Thus, filtering of signals is an important part of pre-processing [3]. The goal is to increase the signal to noise ratio [4], to make feature extraction and pattern recognition easy. Different artifacts have different range of frequencies. Brain electrical activities have a frequency range of 0.3-40Hz [3]–[5]. Any frequency beyond this range is considered noise. The signals are thus passed through appropriate filters like low-pass filters and high pass filters. These filtered signals are then decimated in some cases to further get stronger output. Images are also obtained from some data acquisition methods. Pre-processing of images involve image enhancement, skull removal, de-noising, removal of text (if any) [6].

Features extraction of the processed signals can be done in time and frequency domain. Some of the feature extraction techniques as described in [7] are, Fast Fourier Transform method; Wavelet

Transform method which includes Continuous Wavelet Transform method and Discrete Wavelet Transform; Eigenvector methods which includes Pisarenko's method, Minimum norm method, Music method; Time-frequency distribution methods; Autoregressive methods which includes Yule-Walker method and Burg's method.

In Fast Fourier transform method, data sequence is obtained by power spectral density estimation of signals. It is suitable for narrowband, stationary signals. Although the speed of this method is better [7], it has some disadvantages. It cannot be employed for short signals; it cannot reveal localized spikes and has large noise sensitivity [7]. Wavelet Transform methods are a range of time-frequency domain methods employing spectral estimation technique, where any functional data can be represented as an infinite series of wavelets. This method is suitable for transient and stationary signals [7]. Eigenvector methods calculate signal's frequency and power based on Eigen decomposition. This method is suitable if the signal is still noisy [7]. Feature extracted from Time-frequency distribution methods gives the values of length, frequency and energy of principal track [7]. As described in [7], the advantage of this method is the feasibility of examining great continuous segments of signals. The drawbacks of this method include its slow speed, dependency of extracted features on each other and requirement of an additional pre-processing step known as window process. Autoregressive methods employ parametric approach for feature extraction to calculate power spectral density [7]. In this method, coefficients are calculated which are later used as parameters in linear machine learning models. This method is suitable for signal with sharp signals [7]. Though this method yields good frequency resolution, it is susceptible to biases. Authors in [8] have grouped feature extraction methods in three categories namely, Signal based feature extraction methods, Selection based feature extraction methods and Feature extraction methods for applications. A feature vector is usually formed of the extracted features which can be further used by training algorithms. Feature extraction has to be applied to signals in order to get feature vectors. This is not the case if the data is of type image. Images are directly fed to machine learning or deep learning algorithms after pre-processing. Next, I study how these features obtained from various techniques are utilized in different algorithms.

#### **4. TRADITIONAL MACHINE LEARNING TECHNIQUES**

Machine learning methods in BCI are mainly used to classify various types of signals [3]. Usually the signals that are generated during certain triggers are to be detected at certain times so that they can be used to operate external devices. One such example of a trigger would be invoked by using P300 speller[10] which helps disabled subjects to spell words using brain. A grid of characters is shown to the subject and she looks at the target character. Different rows and columns are intensified in the process and when both the intensified row and column contains the target character, P300 wave is generated by the brain[10]. The task of most of the classifiers is to identify these waves by classifying the signals into P300 and non-P300 signals. In [3], various machine learning methods employed to differentiate EEG segments from different waves say for example, P300 event related potential are described. P300 wave is generated approximately 300ms after exposure to infrequent stimulus. Thus, the feature vector obtained from the pre-processing and filtering stage as mentioned before is fed to classifiers. In some cases, Principal Component Analysis is performed on the feature vectors to reduce the dimensionality before applying classification. Let us first consider the linear classifiers considered in the papers [3], [9], [10]. Linear classifiers determine the parameters and the bias which are used on test data for prediction.

Support vector machines (SVM) are used in pattern recognition in BCI research because it can process high dimensional data. Different kernels like Linear kernel, polynomial kernel, Radial Basis function (RBF) kernel are used in SVM to generate different algorithms. In some research

work, the output signals from SVM are averaged [3] over sequences. In some cases, the output is averaged over sequences and classification score. This approach where classifier output are averaged, help to reduce subject variance [3].

Fisher's Linear Discriminant Analysis (FLDA) is also one of the widely used pattern recognition algorithms. As clearly described in [11], FLDA aims to find a vector that projects all the sample data into a new vector space where the sample data can be distinctly classified into two classes in case of binary classification.

Bayesian Linear Discriminant Analysis (BLDA) is a Bayesian version of Fisher's Linear Discriminant Analysis[3]. To prevent overfitting, the Bayesian Linear Discriminant Analysis makes use of regularization. The degree of regularization can be estimated quickly through Bayesian analysis.

Kernel Fisher Discriminant (KFD) is a non-linear discriminant of Fisher Linear Discriminant. Linear algorithm can be mapped to high dimensional feature space with the help of Kernels [9]. Using this technique, Fisher Linear Discriminant is generalized to its non-linear form.

Random Forest Classifier method involves multiple decision trees. Each decision tree classifies a sample to a particular class. Random Forest classifier then aggregates the results obtained from all of its decision trees and makes a classification decision[10].

Currently SVM, variants of Fisher Linear Discriminant and Random Forest classifier are amongst the widely used classification methods to classify EEG signals. For classifying EEG signals for P300 wave detection, SVM is the most commonly used method [10]. It is observed that experiments performed in [3] that performance of BLDA outperforms SVM. According to [3] Bayesian Linear Discriminant Analysis performs best amongst SVM and the variations of Fisher's Linear Discriminant. This claim is supported by authors in [9], to some extent, where they performed experiments using SVM and Fisher's Linear Discriminant along with PCA for feature reduction. Instead of backing a specific Bayesian version of Fisher's Linear Discriminant, they have proposed their views for Fisher's Linear Discriminant. They proposed SVM to be the best classification method, based on the winner of BCI 2003 and BCI 2005 who used SVM, and the author's work, where SVM shows good results over Fisher's Linear Discriminant. Mirghasemi et.al.[9] demonstrated that Fisher's Linear Discriminant can achieve better performance, given that feature reduction technique Principal Component Analysis be applied before classification. Though both of these papers haven't taken into consideration the effect of Random Forest Classifier. A study conducted in [10] shows that Random Forest classifier performs slightly better than SVM. Possible reasons for better accuracy of Random forest, as mentioned in [10] is because Random Forest classifiers are less sensitive to outliers in training data.

Thus, it can be said that although a lot of methods have proved to perform better and yield better accuracies, SVM still (currently) remains widely used classification technique to classify brain signals.

Next, I investigate the Deep Learning methods currently employed under BCI.

## **5. DEEP LEARNING TECHNIQUES**

Various deep learning methods are employed to classify brain signals and images. Artificial Neural Networks (ANN)[22], Recurring Neural networks (RNN)[12], Convolutional Neural networks (CNN)[12], Deep Belief Networks (DBN)[12], Generative Adversarial Networks(GAN)[14], Variational Autoencoders(VAE)[14], Long Short Term Memory[16], Echo

State Network[12] are Deep Learning techniques currently being employed in the field of Brain Computer Interface.

Artificial Neural networks consists of layers of multiple input and output nodes along with one or more layers of hidden nodes. They are statistical data modelling tools which are non-linear in nature. They are vaguely inspired by the neurons and the neural network of brain. Convolutional Neural Networks are a type of Artificial Neural Networks which consists of Convolutional layers and pooling layers. Recurrent Neural networks are also a type of Artificial Neural Networks where connections between neurons make a directed cycle. Thus, RNNs take into consideration the state of previous neurons along with its own input. Deep Neural Network is any Artificial Neural network consisting many hidden layers, thus making the network more deep. Restricted Boltzmann Machines are a class of Artificial Neural Networks that can learn probability distribution of its input. Deep Belief Networks are formed by stacking multiple Restricted Boltzmann Machines. Generative Adversarial Networks are a class of generative models which consists of two competing Neural Networks. Autoencoders are also of Artificial Neural Networks which aims to learn a function representation for a set of data. Variational autoencoders along with learning a function that represents data, also learns the probability distribution representing the data. Long Short-Term Memory (LSTM) is a type of Recurrent Neural Network which has nodes called memory units which can retain information for long periods of time. Echo State Networks are a type of Recurrent Neural Network which has sparsely connected hidden layers.

Authors in [12] have presented a detailed overview of different deep learning methods employed in this field. Deep Belief Networks, Convolutional Neural Networks and Recurrent Neural Networks are used to learn data representation of time series EEG data [12]. Autoencoders are trained to focus on stable features [12] and a similarity constraint is used to tell features from different apart. Deep Belief Networks are also used to classify data based on Motor Imagery tasks [12]. In [13], three models of Artificial Neural networks, viz., Multi-layer Perceptron, Elman Recurrent Neural Network and Time-dependent Recurrent Neural Networks are used. They conclude that though EEG data is a sequence of vectors, applying Recurrent Neural Networks to EEG data is not straightforward [13].

Besides classification, Neural Networks can also be used to generate data. Generative Adversarial Networks and Variational Autoencoders are generative models which can employed for this purpose. Generative Adversarial Networks are used in BCI to generate EEG signals as seen in [14] and [15]. They outperform Variational autoencoders [14] in data generation. Generative Adversarial Networks are also used to interpret the resemblance between original EEG data and EEG patterns that deep convolutional network has learned [14]. Thus GANs can be used to generate artificial EEG data which can be used for data restoration and data augmentation [14]. In [15], a semi-supervised variation of Generative Adversarial Networks are used which employs Convolutional layers for data generation.

EEG data is usually high dimensional and multichannel. Thus, Deep Learning techniques can also be applied to raw data to extract features. Some authors like in [16]–[18] have demonstrated the same. In [16] Long Short-Term Memory and Convolutional Neural Networks are used for feature extraction in time domain. In [17] Convolutional Deep Belief Networks are used to extract features. In [18], a Deep Neural Network is used to extract high-level representation of raw EEG data. They also use Principal Component Analysis to extract components which are most relevant, to address the problem of overfitting.

Various Ensemble Learning methods are employed to obtain better results. Ensemble learning is a method where various models and classifiers are combined to solve a problem. In [19], a deep neural network is proposed for motor imagery classification. It is a combination of Convolutional

Neural Networks and stacked Autoencoders. In [20], 2nd BCI data competition dataset is classified using a model which is combination of MELM (Multi-layer Extreme Learning Machine) and KELM (Extreme Learning Machine with Kernel). In [21], two emotional categories are classified based on EEG signals. They trained a Deep Belief Network with differential entropy and Hidden Markov Models to capture emotional stage switching.

Transfer technique is also employed by some researchers in their model. Transfer Learning is a method where a pretrained network is used. While training such networks, only the output layers are trained and the pretrained network is kept frozen. After training the output layer, all the layers along with the pretrained network are activated. One such technique is demonstrated in [22] where pre-trained networks renet101[22] and vgg19[22] are used for seizure detection in EEG signals.

According to [12], deep learning techniques like Deep Belief Networks, Convolutional Neural Networks and Echo state networks learn EEG data representation better than traditional machine learning techniques. To learn invariant features of brain activity, deeper level of data representation is more favourable [12]. This claim is also supported by authors in [23]. Their experimental findings in [23] indicate that deep learning on average performs better than classic machine learning techniques. The same is also claimed in [24] where additionally they have shown that using deep learning techniques reduces the dependency on feature extraction. Deep Neural Networks are known to work best with huge datasets. Thus, Hennrich et.al. [25] recommends using regularized classifiers like LDA over deep neural networks if the training data is less.

In summary, Deep Belief Networks, Convolutional Neural Networks and Recurrent Neural Networks are widely used for classification purposes along with ensemble models. Generative Adversarial Networks and Variational Autoencoders are used mainly for data generation. Thus, Deep Learning techniques, though widely employed to represent and classify brain data, are used in recent studies to generate data for data augmentation and for feature extraction. It is claimed and proved by a lot of studies that deep learning methods outperform classic machine learning techniques for classification if the dataset is large.

## **6. CONCLUSION**

Thus, this paper introduces in brief the field of Brain computer interface. It summarizes different methods of data acquisition to obtain brain data. A number of feature extraction techniques are described to obtain features in time and frequency domain. Furthermore, a number of Machine Learning methods are introduced which are currently being employed in this field for classification of data. A brief comparison is provided and widely used techniques are described. Finally, a study is presented on various Deep Learning methods in this field. After defining each technique, their applications are described. It can be inferred that deep learning methods are not only superior in performance than machine learning techniques but are also more prevalent. Deep learning techniques are also applied to a wide range of problems apart from classification. Thus, deep learning techniques are more prevalent in this field as compared to machine learning techniques, given their performance and their use in wide applications.

## REFERENCES

- [1] F. Lotte and I. B. Sud-ouest, "Brain-Computer Interfaces : Beyond Medical Applications," IEEE, pp. 26–34, 2012.
- [2] N. Elsayed, Z. Saad, and M. Bayoumi, "Brain Computer Interface: EEG Signal Preprocessing Issues and Solutions," *Int. J. Comput. Appl.*, vol. 169, no. 3, pp. 12–16, 2017.
- [3] A. E. Selim, M. A. Wahed, and V. M. Kadah, "Machine learning methodologies in P300 speller brain-computer interface systems," *Natl. Radio Sci. Conf. NRSC, Proc.*, 2009.
- [4] H. Xu, J. Lou, R. Su, and E. Zhang, "Feature extraction and classification of EEG for imaging left-right hands movement," *Proc. - 2009 2nd IEEE Int. Conf. Comput. Sci. Inf. Technol. ICCSIT 2009*, pp. 56–59, 2009.
- [5] E. Estrada, H. Nazeran, P. Nava, K. Behbehani, J. Burk, and E. Lucas, "EEG feature extraction for classification of sleep stages," *Annu. Int. Conf. IEEE Eng. Med. Biol. - Proc.*, vol. 26 I, pp. 196–199, 2004.
- [6] B. D. Rao and M. M. Goswami, "A comprehensive study of features used for brain tumor detection and segmentation from Mr images," *2017 Innov. Power Adv. Comput. Technol. i-PACT 2017*, vol. 2017-January, pp. 1–6, 2017.
- [7] A. S. Al-Fahoum and A. A. Al-Fraihat, "Methods of EEG Signal Features Extraction Using Linear Analysis in Frequency and Time-Frequency Domains," *ISRN Neurosci.*, vol. 2014, pp. 1–7, 2014.
- [8] J. C. Mohammad Rahman, Wanli Ma, Dat Tran, "A Comprehensive Survey of the Feature Extraction Methods in the EEG Research," in *Conference: Proceedings of the 12th international conference on Algorithms and Architectures for Parallel Processing - Volume Part II*, 2012.
- [9] H. Mirghasemi, R. Fazel-Rezai, and M. B. Shamsollahi, "Analysis of P300 classifiers in brain computer interface speller," *Annu. Int. Conf. IEEE Eng. Med. Biol. - Proc.*, pp. 6205–6208, 2006.
- [10] F. Akram, S. M. Han, and T. S. Kim, "An efficient word typing P300-BCI system using a modified T9 interface and random forest classifier," *Comput. Biol. Med.*, vol. 56, pp. 30–36, 2015.
- [11] Y. J. Chiou et al., "Unsupervised classification for volume-based magnetic resonance brain images," *Proc. - 2014 Int. Symp. Comput. Consum. Control. IS3C 2014*, no. 5, pp. 621–624, 2014.
- [12] L. Bozhkov and P. Georgieva, "Overview of Deep Learning Architectures for EEG-based Brain Imaging," *Proc. Int. Jt. Conf. Neural Networks*, vol. 2018-July, pp. 1–7, 2018.
- [13] A. S. Greaves, "Classification of EEG with Recurrent Neural Networks," pp. 1–5, 2011.
- [14] F. Fahimi, Z. Zhang, W. B. Goh, K. K. Ang, and C. Guan, "Towards EEG generation using gans for bci applications," *2019 IEEE EMBS Int. Conf. Biomed. Heal. Informatics, BHI 2019 - Proc.*, pp. 1–4, 2019.
- [15] W. Ko, E. Jeon, J. Lee, and H. Il Suk, "Semi-Supervised Deep Adversarial Learning for Brain-Computer Interface," *7th Int. Winter Conf. Brain-Computer Interface, BCI 2019*, pp. 1–4, 2019.
- [16] P. Lu, N. Gao, Z. Lu, J. Yang, O. Bai, and Q. Li, "Combined CNN and LSTM for Motor Imagery Classification," *Proc. - 2019 12th Int. Congr. Image Signal Process. Biomed. Eng. Informatics, CISP-BMEI 2019*, pp. 1–6, 2019.
- [17] Y. Ren and Y. Wu, "Convolutional deep belief networks for feature extraction of EEG signal," *Proc. Int. Jt. Conf. Neural Networks*, pp. 2850–2853, 2014.

- [18] S. Jirayucharoensak, S. Pan-Ngum, and P. Israsena, "EEG-Based Emotion Recognition Using Deep Learning Network with Principal Component Based Covariate Shift Adaptation," *Sci. World J.*, vol. 2014, 2014.
- [19] Y. R. Tabar and U. Halici, "A novel deep learning approach for classification of EEG motor imagery signals," *J. Neural Eng.*, vol. 14, no. 1, 2017.
- [20] S. Ding, N. Zhang, X. Xu, L. Guo, and J. Zhang, "Deep Extreme Learning Machine and Its Application in EEG Classification," *Math. Probl. Eng.*, vol. 2015, 2015.
- [21] W. L. Zheng, J. Y. Zhu, Y. Peng, and B. L. Lu, "EEG-based emotion classification using deep belief networks," *Proc. - IEEE Int. Conf. Multimed. Expo*, vol. 2014-September, no. September, pp. 1–6, 2014.
- [22] A. Agrawal, G. C. Jana, and P. Gupta, "A deep transfer learning approach for seizure detection using RGB features of epileptic electroencephalogram signals," *Proc. Int. Conf. Cloud Comput. Technol. Sci. CloudCom*, vol. 2019-December, pp. 367–373, 2019.
- [23] N. Nagabushan, T. Fisher, G. Malaty, M. Witcher, and S. Vijayan, "A comparative study of motor imagery based BCI classifiers on EEG and iEEG data," *Glob. 2019 - 7th IEEE Glob. Conf. Signal Inf. Process. Proc.*, 2019.
- [24] J. Thomas, T. Maszczyk, N. Sinha, T. Kluge, and J. Dauwels, "Deep Learning-based Classification for Brain-Computer Interfaces," pp. 234–239, 2017.
- [25] J. Hennrich, C. Herff, D. Heger, and T. Schultz, "Investigating deep learning for fNIRS based BCI," *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, vol. 2015-November, pp. 2844–2847, 2015.