JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

EEG SIGNAL ANALYSIS FOR EPILEPSY DISEASE USING MACHINE LEARNING TECHNIQUES

[1]K.Navya Manjeera, [2]P.Sekhar, [3]P.Harshavardhan, [4]P.Sai Ganesh, [5]M.Jayanth, [6]V.Jagan Naveen

[12345]B.Tech Students, [6]Professor, Department of Electronics and Communication Engineering,
GMRIT College of Engineering, Rajam, Andhra Pradesh, India

Abstract: Epilepsy is recognized to be one among the most critical and central nervous system (neurological) disorder affecting the human brain. Epilepsy can be detected with the help of Electroencephalography (EEG), EEG is an effective technique which is used to monitor the brain activity. EEG also used for diagnosing epilepsy. By analyzing the EEG raw data epilepsy can be detected at early stages. In this paper, we presented a technique for detecting the disease using EEG raw data. The proposed model is made using a neural network technique called artificial neural network (ANN).The model will be trained using the larger dataset of patient diseases. Finally, the model will be tested against random samples to obtain the results. The results obtained 92.6% of maximum accuracy using ANN.

IndexTerms -Epilepsy, Electroencephalography, Artificial Neural Network, Machine Learning Signal Analysis, Medical Applications.

I. INTRODUCTION

One of the most prevalent neurological conditions affecting people worldwide is recognized as epilepsy. According to the WHO, it is the second most prevalent neurological condition that affects millions of people, only after stroke. A seizure occurs in one out of every 100 people at some point in their lifetime. Most epilepsy cases start at an early stage or teenage and some cases start in old age. Seizure is not the same disorder as epilepsy or can say that all seizures are not epilepsy fits. People affected with epilepsy will suffer from sudden seizures. In this period, it is difficult for them to protect themselves. It happens when a normal neuronal network suddenly turns into a hyper-excitable network. This process is called as epileptogenic, which affects mainly the cerebral cortex. The seizures characteristics vary and depend upon the disturbance first starts in the brain, and far it spreads. The symptoms that occur are as follows: disturbances of movement, loss of consciousness, and the problems in hearing, vision, mood, or other cognitive functions.
In order to identify an epileptic seizure, electroencephalography (EEG) is essential. To detect the ionic currents moving within the brain, electrodes are positioned around the scalp, and the voltage variations and differences between them are measured.

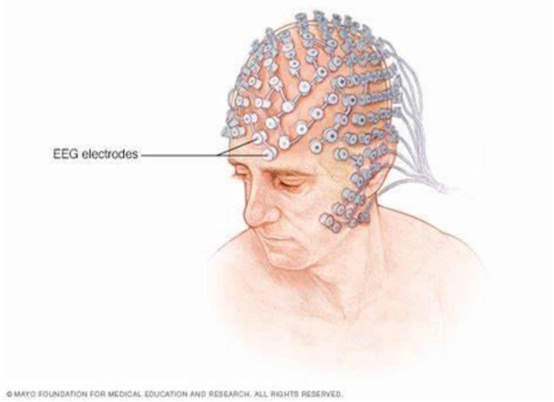


Fig 1: EEG Electrodes

The detection with EEG requires immediate examination by a physician along with the substantial amount of effort and time. It is difficult to analyze the recordings and detect the epilepsy is complicated. Therefore, automation, and a computer-aided method is urgently needed for the diagnosis of epilepsy.

Basically, diagnosis is the principal purpose of EEG signals. All forms of brain problems, such as epilepsy, tremors, concussions, strokes, and sleep disturbances are treated by it. In more recent EEG apps, machine learning is used as an analysis technique. Numerous studies have focused in particular on using machine learning to diagnose sleep problems and epileptic seizures. EEG signals are used in games that control and modify objects while tracking brain activity.

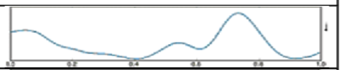
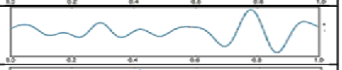
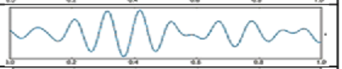
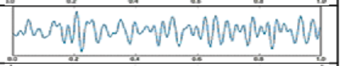
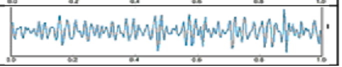
Waves	Frequency bands (Hz)	Behaviour Trait	Signal Waveform
Delta	0.3 – 4	Deep sleep	
Theta	4 – 8	Deep Meditation	
Alpha	8 – 13	Eyes closed, awake	
Beta	13 – 30	Eyes opened, thinking	
Gamma	30 and above	Unifying consciousness	

Fig 2: Frequency Bands of EEG Waveforms

The Electroencephalograph measures the waves of various frequencies within the brain. If the frequencies are excessive, imperfect, or difficult to access, the mental performance can suffer. The raw data of EEG has usually described in terms of frequency bands, those are: Gamma waves which are greater than 30Hz, these waves are recognized as the fastest brain waves. Gamma waves mainly occur when you are conscious and highly alert state**.** Beta waves are in the range of 13-30Hz, these waves occurred in the state of stress and a phenomenon called ‘Alpha blocking’ and they do not occur in humans until three years of age**.** Alpha waves are in the range of 8-12 Hz, Alpha waves occurred in awake adults while resting with the eyes closed. Theta waves are in the range of 4-8Hz, Theta waves occurred during the sleeping or dreaming. Delta waves are in the range less than 4 Hz, these waves occurred in the deep sleep and occurred in some abnormal process.

EEG offers a cheap and efficient approach to examine brain function, among other benefits. There is no risk of radiation from the object.

Related works:

[1]They proposed a model for detection of epilepsy by using a machine learning technique. The KNN and SVM are using to detect the seizure. Data with signals are recorded with a sampling rate of 173.6 segments. The data is recorded from ten people in them five are from epilepsy patients during seizure and another five are healthy people. Model was made under using the extraction of feature of each wave in frequency domain, wavelet Transform, time domain. Feature selection done with SFFS, T- test. By considering the Sensitivity, Specificity, and accuracy of the SVM model has been good over KNN.

[2]In this paper, they applied the algorithms such as SVM, RF, KNN, Decision Tree, Artificial Neural Network (ANN) using the PCA reduction technique to detect epilepsy. They analyzed these algorithms by the performance using PCA and without PCA. They divided dataset by setting 75% to training and 25% for testing. This model using Random Forest detected higher accuracy of 97% with low Computational time compared to the other algorithms.

[3]In this paper, they proposed the improved RBF model to detect epilepsy using EEG. To classify signals of epileptic EEG accurately, they had done research to analyze features from non-linear and linear interrelations and they put those as input to the RBF model to get effective features. They also introduced the OAO method to detect the classification for epilepsy by computer.

[4]They proposed the model by using the SVM and KNN to detect epilepsy. They also compared these algorithms by considering the accuracy and latency of these models respectively. They used the dataset from CHBMIT Scalp EEG Database. They used the sampling frequency of 256 Hz, which means that 256 data points per second were recorded. They divided into three second segments. They concluded that KNN performed better than SVM with 78% accuracy respectively.

[5]In this paper, they proposed automatic detection for epilepsy using various time-frequency approaches. They developed Machine learning technique Novel RF-GSO as, automatic epilepsy seizure detection. This is based on PCA analysis along with time-frequency analysis. The novel model has its limitations. The detection will be affected, if the noise is too high. Before using this model, it is necessary to do pre-processing to remove the noise in EEG. This model can detect and decisions more effectively and accurately.

[6] They suggested using machine learning to analyze EEG data in order to identify the beginning of an epileptic seizure. They considered and evaluated the SVM method which they worked on the CHB-MIT database. They trained this method for 916 hours on data collected from 24 patients. Their model detected the percentage of 96% of 173 test seizures, and delay of 3 seconds for median detection.

[7] To identify epileptic seizures, they suggested using a CNN imaged-EEG model representation. They made advantage of the data set gathered from the EPILEPSIAE study and the CHB-MIT database, which both contain intracranial and scalp recordings. The

proposed method got precision of 62.7%, recall of 58.3% for a CHB-MIT database, and performances for the EPILEPSIAE data are low comparatively. There proposed model got accuracy of 99.3 and 98.0 respectively.

[8]They proposed detection models for the epileptic seizure event and onset detection. They used a Bonn University EEG database and a CHB-MIT database for onset. The data is recorded with the Daubechies wavelet features which are computed at the third and fourth level. The features are maximum, minimum, standard deviation, entropy, mean, and energy. They achieved accuracy of 84.2%.Within the case of onset detection, and a sensitivity of 98.5% has been achieved at 1.76 seconds with a mean latency.

[9]They proposed a Discrete Cosine Transformation (DCT) type II algorithm for detecting seizure, which extracts features of energy from sub-bands by transforming the signal into frequency-domain. They suggested a technique to choose the ideal subset from a total of 23 channels by taking into account their maximum variation. They employed KNN to identify frames that were interictal or at ictal distance for Euclidean distance. They had achieved 93.12 of F1-score, and a False-Positive Rate average of 0.07 where they tested on CHB-MIT dataset which includes data of 21 patients.

II. METHODOLOGY

Dataset:

The dataset used here is taken from Kaggle**.** This dataset is about confused student EEG brainwave data. The features of the dataset are as follows:

1. Subject Id
2. Video Id
3. Attention
4. Mediation
5. Raw
6. Delta
7. Theta
8. Alpha 1
9. Alpha 2
10. Beta 1
11. Beta 2
12. Gamma 1
13. Gamma 2
14. Predefined labeln
15. User-defined labeln
16. Gender
17. Age
18. Ethnicity

Below is a representation of the model's block diagram.

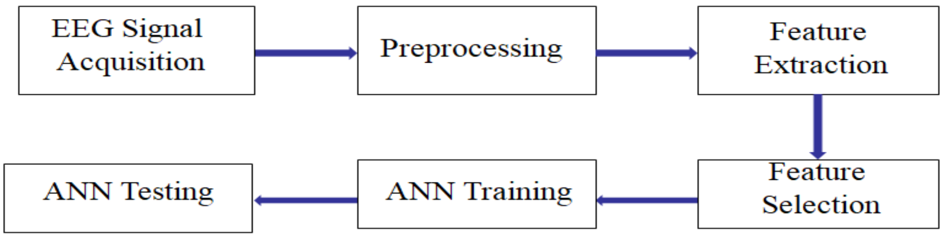


Fig 3: Block Diagram

EEG signal acquisition is the initial stage. This is essentially unprocessed raw data. Pre-processing, which occurs in the second phase, can help the model's accuracy by removing noise and other outliers from the dataset. The third phase is feature extraction. Based on the requirements of the machine learning algorithms, the conversion of raw data into inputs takes place. The fourth phase is feature selection which is the selection of required parameters. The fifth phase is ANN training. Here the model will be trained with the help of the training dataset. The last phase is ANN testing. Testing set consists of the unknown data which learns on the training set to evaluate the performance.

Implementation of machine learning algorithm:

A computational model called Artificial Neural Networks (ANN) imitates how nerve cells function in the human brain. ANN employs learning algorithms that, in a sense, "learn" as they receive new data and can independently make corrections. The benefit of utilizing ANN is that it can learn and model complex, non-linear interactions. After learning from the original inputs and their associations, it may also generalize and predict on previously unexplored data. Due to the non-linear and complicated nature of many relationships between inputs and outputs in real life, these characteristics of the ANN make it extremely significant. In this work a new model is proposed to identify the epilepsy disease from the given data set.

The following is the suggested architecture's block diagram.

```
Model: "sequential_1"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_9 (Dense)             (None, 64)                960

 batch_normalization_8 (Bat  (None, 64)                256
 chNormalization)

 dropout_8 (Dropout)         (None, 64)                0

 dense_10 (Dense)            (None, 124)               8060

 batch_normalization_9 (Bat  (None, 124)               496
 chNormalization)

 dropout_9 (Dropout)         (None, 124)               0

 dense_11 (Dense)            (None, 248)               31000

 batch_normalization_10 (Bat (None, 248)               992
 chNormalization)

 dropout_10 (Dropout)        (None, 248)               0

 dense_12 (Dense)            (None, 512)               127488

 batch_normalization_11 (Bat (None, 512)               2048
 chNormalization)

 dropout_11 (Dropout)        (None, 512)               0

 dense_13 (Dense)            (None, 664)               340632

 batch_normalization_12 (Bat (None, 664)               2656
 chNormalization)

 dropout_12 (Dropout)        (None, 664)               0

 dense_14 (Dense)            (None, 512)               340480

 batch_normalization_13 (Bat (None, 512)               2048
 chNormalization)

 dropout_13 (Dropout)        (None, 512)               0

 dense_15 (Dense)            (None, 264)               135432

 batch_normalization_14 (Bat (None, 264)               1056
 chNormalization)
```

```
 dropout_14 (Dropout)        (None, 264)              0

dense_16 (Dense)            (None, 124)              32860

 batch_normalization_15 (Bat (None, 124)              496
chNormalization)

 dropout_15 (Dropout)        (None, 124)              0

 dense_17 (Dense)            (None, 2)                250

=================================================================
Total params: 1,027,210
Trainable params: 1,022,186
Non-trainable params: 5,024
```

____

Fig 4: Proposed model architecture

After passing it through the batch normalization layer, the ReLu activation function is used at the output end. To prevent the overfitting in the model, a dropout rate of 0.27 was used. This model has a total of 1,027,210 parameters out of which 1,022,186 are trainable and 5,024 are non-trainable parameters. This model is compiled using Adamax optimizer with a learning rate of 0.001 and loss is set to binary cross entropy.

Evaluation Criteria:

To assess the effectiveness of the model single metric called accuracy is not sufficient. Apart from accuracy other evaluation metrics such as precision and recall should be calculated. The three metrics are as follows:

1. Accuracy: It is defined as the ratio of the correct predictions over the total predictions.

$$\text{Accuracy} = \frac{(TP+TN)}{(TP+TN+FP+FN)}$$

2. Precision: It is defined as the ratio of the correct classified positive predictions to the total number of classified positive predictions

$$\text{Precision} = \frac{(TP)}{(TP+FP)}$$

3. Recall: The percentage of Positive samples that were accurately labelled as Positive relative to all Positive samples is how the recall is calculated.

$$\text{Recall} = \frac{(TP)}{(TP+FN)}$$

Here TP, TN, FP, and FN are the true-positive, true-negative, false-positive, and false-negative samples.

III. RESULTS AND DISCUSSION

In this section, we will discuss about the effectiveness of the suggested model. Along with the other three pre-trained models, this model is also tested for epilepsy detection using the same dataset.

Machine Learning algorithms	Accuracy (%)
SVM	67.2
Bi-LSTM	73.3
CF-Bi-LSTM	75.0
This work	**92.6**

Table 1: The performance of the three suggested pre-trained models.

A. PERFORMANCE OF THIS WORK

Three pre-trained models are being used on the same dataset to compare the performance of the proposed model. These models are (i) SVM (ii) Bi-LSTM and (iii) CT-Bi-LSTM. The performance of our model is then compared with these pre-trained models in Table 2 in terms of accuracy. Thus, our proposed model ranks first and achieves an accuracy of 92.6% and it outperformed the other three pre-trained models.

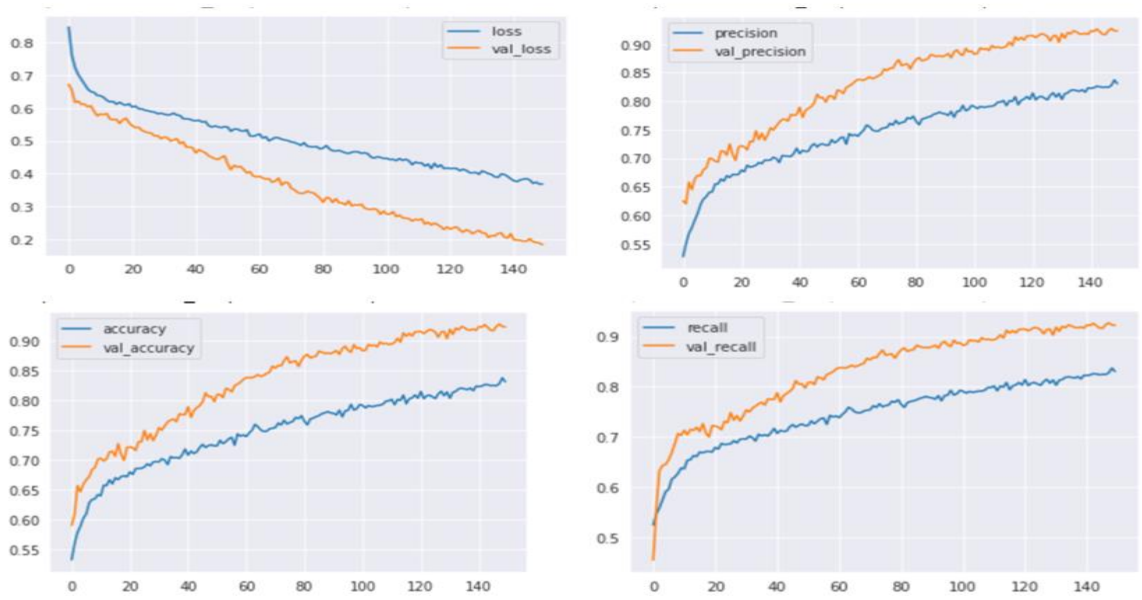


Fig 5: Loss, accuracy, precision and recall.

Fig.5 shows the loss, accuracy, precision and recall, validation loss, validation accuracy, validation precision and validation recall graphs of the model respectively. The X-axis represents the number of epochs, and the Y-axis represents the value of the loss, accuracy, precision, recall, validation loss, validation accuracy, validation precision and validation recall.

B. COMPARISON

We observed that our model has achieved high accuracy, precision and recall compared to the previous works of [2] and [6]. Our model has achieved less accuracy when compared to the [2][6]. But when compared to [2][6,] our model has fared better in terms of recall and precision. [2] used machine learning-based algorithm called Random Forest to detect epilepsy with accuracy of 97%. [4] used machine learning-based algorithm called as KNN to detect epilepsy with accuracy of 78%. [6] used Machine learning-based algorithm called as SVM to detect epilepsy with accuracy of 96%.

Approach	Dataset	Accuracy(%)	Precision(%)	Recall (%)
Artificial Neural Networks	Kaggle	92.6	92.6	92.6

Table 2: New data prediction

IV. CONCLUSION AND FUTURE SCOPE

In this paper, we introduced a machine learning-based approach for detecting epilepsy. Initially, the data of confused levels of students are taken, and they are rearranged. To make model much accurate we pre-processed the data to remove the noise and other outliers, and then the data is feed to the model. We developed this model by using activation function, loss function, different layers, and algorithms to reduce the estimation cost without affecting the accuracy of the model. By comparing our model with the pre-trained models, i.e., SVM, Bi-LSTM and CF-Bi-LSTM this model achieved comparatively better performance. According to the experimental results the proposed approach has accomplished an accuracy of 92.6%, precision of 92.6% and recall of 92.6%. Therefore, we conclude that our model outperformed all the other models for the prediction of epileptic detection.

In the future, we must improve the pre-processing phase of EEG signals to get an improved accuracy of epileptic prediction.

V. ACKNOWLEDGMENT

I would like to thank our guide Dr. V. Jagan Naveen, Professor and HOD, Dept of ECE of wholehearted and valuable guidance throughout the project, and I would like to sincerely thank our Principal, all the staff members, and teammates for the direct and indirect support in helping us in the completion of the project.

VI. REFERENCES

[1] Savadkoohi M, Oladunni T, Thompson L. A machine learning approach to epileptic seizure prediction using Electroencephalogram (EEG) Signal. Biocybernetics and Biomedical Engineering. 2020 Jul 1; 40(3):1328-41.

[2] Nahzat S, Yaganoglu. Classification of Epileptic Seizure Dataset Using Different Machine Learning Algorithms and PCA Feature Reduction Technique. Journal of Investigations on Engineering and Technology. 2021; 4(2):47-60.

[3] Zhou D, Li X. Epilepsy EEG signal classification algorithm based on improved RBF. Frontiers in Neuroscience. 2020 Jun 23; 14:606.

[4] Kang Lövgren S, Rosquist C. Machine Learning Methods for EEG-based Epileptic Seizure Detection.

[5] Wang X, Gong G, Li N, Qiu S. Detection analysis of epileptic EEG using a novel random forest model combined with grid search optimization. Frontiers in human neuroscience. 2019 Feb 21; 13:52.

[6] Shoeb AH, Guttag JV. Application of machine learning to epileptic seizure detection. InICML 2010 Jan 1.

[7] Gómez C, Arbeláez P, Navarrete M, Alvarado-Rojas C, Le Van Quyen M, Valderrama M. Automatic seizure detection based on imaged-EEG signals through fully convolutional networks. Scientific reports. 2020 Dec 11; 10(1):1-3.

[8] Ahammad N, Fathima T, Joseph P. Detection of epileptic seizure event and onset using EEG. BioMed research international. 2014 Jan 29; 2014.

[9] Jumaah MA, Shihab AI, Farhan AA. Epileptic Seizures Detection Using DCT-II and KNN Classifier in Long-Term EEG Signals. Iraqi Journal of Science. 2020 Oct 28:2687-94.

[10] Hosseini MP, Hosseini A, Ahi K. A review on machine learning for EEG signal processing in bioengineering. IEEE reviews in biomedical engineering. 2020 Jan 28; 14:204-18.

[11] Subasi A. Application of adaptive neuro-fuzzy inference system for epileptic seizure detection using wavelet feature extraction. Computers in biology and medicine. 2007 Feb 1; 37(2):227-44.

[12] Prabhakar SK, Rajaguru H. PCA and K-means clustering for classification of epilepsy risk levels from EEG signals—a comparitive study between them. In2015 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS) 2015 Nov 28 (pp. 83-86). IEEE.

[13] Günay M, Ensari T. EEG signal analysis of patients with epilepsy disorder using machine learning techniques. In2018 Electric Electronics, Computer Science, Biomedical Engineerings' Meeting (EBBT) 2018 Apr 18 (pp. 1-4). IEEE.

[14] Siddiqui MK, Morales-Menendez R, Huang X, Hussain N. A review of epileptic seizure detection using machine learning classifiers. Brain informatics. 2020 Dec; 7(1):1-8.

[15] Subasi A, Kevric J, Abdullah Canbaz M. Epileptic seizure detection using hybrid machine learning methods. Neural Computing and Applications. 2019 Jan; 31(1):317-25.

[16] Shoeibi A, Khodatars M, Ghassemi N, Jafari M, Moridian P, Alizadehsani R, Panahiazar M, Khozeimeh F, Zare A, Hosseini-Nejad H, Khosravi A. Epileptic seizures detection using deep learning techniques: a review. International Journal of Environmental Research and Public Health. 2021 May 27; 18(11):5780.

[17] Abbasi B, Goldenholz DM. Machine learning applications in epilepsy. Epilepsia. 2019 Oct; 60(10):2037-47.

[18] Suguna P, Kirubagari B, Umamaheswari R. An Analysis of Epileptic Seizure Detection and Classification Using Machine Learning-Based Artificial Neural Network. InInventive Systems and Control 2021 (pp. 53-59), Springer, Singapore.

[19] Usman SM, Usman M, Fong S. Epileptic seizures prediction using machine learning methods. Computational and mathematical methods in medicine. 2017 Dec 19; 2017.

[20] Chakraborti S, Choudhary A, Singh A, Kumar R, Swetapadma A. A machine learning based method to detect epilepsy. International Journal of Information Technology. 2018 Sep; 10(3):257-63.