PAPER • OPEN ACCESS

Comparative analysis of multiple machine learning algorithms for epileptic seizure prediction

View the article online for updates and enhancements.

Comparative analysis of multiple machine learning algorithms for epileptic seizure prediction

H O Lekshmy[1], Dhanyalaxmi Panickar[2] and Sandhya Harikumar[3]

Dept of CSE, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Amritapuri, Kerala, India.

E-mail: lekshmyho@am.amrita.edu[1], amenp2ari20012@am.students.amrita.edu[2], sandhyaharikumar@am.amrita.edu[3]

Abstract. Epilepsy is a common neurological disease that affects more than 2 percent of the population globally. An imbalance in brain electrical activities causes unpredictable seizures, which eventually leads to epilepsy. Neurostimulators have the power to intervene in advance and avoid the occurrence of seizures. Its efficiency can be increased with the help of heuristics like advanced seizure prediction. Early identification of preictal state will help easy activation of neurostimulator on time. This research concentrates on the performance analysis of various machine learning algorithms on recorded EEG data. Through this study, we aim to find the best model, which can be used to create an ensemble model for better learning. This involves modeling and simulation of classical machine learning technique like Logistic regression, Naive Bayes model, K nearest neighbors Random Forest, and deep learning techniques like an Artificial neural network, Convolutional neural networks, Long short term memory, and Autoencoders. In this analysis, Random Forest and Long Short-Term Memory performed well among all models in terms of sensitivity and specificity.

Index Terms Epilepsy Prediction, Machine Learning, Deep learning, Neural Network,

Classifiers, Seizure Prediction, Random Forest, LSTM

1. Introduction

Over 50 million people around the world suffer because of epilepsy. Elliptical seizures create changes in behavior, muscle tone, or movements. This eventually leads to vigorous shaking, breathing problems, loss of consciousness, and so on. These seizures occur without any warning and thus leaving no time for patients to prepare. The chance of untimely death of people with epilepsy is nearly three times greater than healthy people. Patients suffering from epilepsy avoid many social activities of their life due to fear of stigma and discrimination. Even with the present advancement in the medical field, we could be able to control only 70% of seizure cases using medicines [1]. The unpredictable nature of this disease causes anxiety and sometimes even physical injuries in patients. With the accurate prediction of upcoming seizures, the anxiety in patients can be reduced. This also provides them with adequate time to put themselves in a

safe position to avoid any injury. Neurostimulators can intervene in advance to avoid the occurrence of seizures [2]. With a proper seizure prediction algorithm, one can easily activate neurostimulators before the seizure. This early prediction will increase the efficiency of the neurostimulator.

Nowadays, Machine learning techniques have been used extensively in many health care applications like anomaly detection in Phonocardiogram [3], Hepatitis disease prediction[4], Breast Cancer Diagnosis [5], and others. A seizure can be detected at the early stages with classical machine learning techniques and deep learning models. The primary method of seizure detection is by analyzing patterns in electroencephalogram (EEG) [6]. EEG data has application in a wide variety of problems like emotion detection [7], EEG-based brain control[8] and so on. The machine learning approach looks through the recordings of EEG and identify the brain activity patterns during seizures. The seizure prediction task makes use of classical machine learning methods like the random forest [9], SVM [10], and so on. But these classical machine learning algorithms are less efficient in high volume EEG data. Variability in data also affects the performance of algorithms. Availability of voluminous data and high-end processors lead to advancement in deep learning techniques. This eventually eased the challenges in seizure predictions.

In this research, we focus on the performance of various machine learning techniques (Logistic regression, Naive Bayes, Random Forest, and K- Nearest neighbor, Artificial neural network, Convolutional neural network, Long short term Memory technique and Auto Encoder) on EEG data to find the best machine learning algorithms that perform well on our datasets. From our analysis, we understand that among classical machine learning techniques, the Random Forest classifier has performed remarkably well in terms of specificity and sensitivity. Among deep learning algorithms, Long Short Term Memory (LSTM) is the best performing model.

2. Literature Survey

Epilepsy detection has been studied extensively since the 1970s. Early, statistical methods and classical machine learning techniques were used for Epilepsy classification. But due to the data complexity, the researchers started applying deep learning to this problem. Since 2016, ample research has been done on it to identify epilepsy using deep learning models. Many deep learning models such as Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Deep Belief Networks (DBN), Autoencoders (AE), Long short-term memory (LSTM), and generative adversarial networks (GAN) have been used [6]. Feature extraction is a key step in the epilepsy detection technique. Various techniques have been used to get relevant features from the EEG. This helps in getting better accuracy and faster computation. For feature extraction, many techniques have been used in recent years which include discrete wavelet transform (DWT)[7] techniques based on Huffman coding [8], Hilbert Transform [9], spectral analysis, and Wavelet Transform [10]. By the 20th century, deep learning models took off. Artificial neural networks along with the sliding windows are utilized for pattern recognition [11]. This was done on the Bonn data set [12]. An accuracy of 93.7% accuracy in detecting normal signals and 95.06% accuracy in detecting seizure signals were observed. Later many new approaches were proposed that make use of more deep learning algorithms. One such approach was to use Convolutional Neural Networks for interictal epileptiform discharge [13]. This is a common symptom of partial epilepsy. This is the first paper that put forward the idea of using automatic feature extraction techniques. But they were focused on the detection of interictal epileptiform discharge rather than predicting seizure. More complex neural networks like Recurrent neural networks were introduced after that. The time-series nature of EEG signals

encouraged researchers to use recurrent neural networks for encephalogram analysis, for the detection of seizure foci [14]. Another approach makes use of a residual network for epilepsy seizure prediction [15]. This model uses ResNet-50 for classifying EEG into 3 classes (ictal, neonatal, and pre-ictal). This classification technique gives an accuracy of 94:98%, but the main drawback was, this model only takes 2D EEG images as input. So, one had to convert 1D EEG data to two 2D EEGs. This incorporated additional computational cost and complexity in the model. Various comparative studies were held to identify the best model. One such research tried to make a comparative analysis of different networks available for seizure detection [16]. This paper covers a comparative study of three different deep learning models i.e., CNN, RNN, and CNN. Training and the test data set in this paper are obtained from totally separate groups of mice. Later, these classifiers were applied to human encephalogram recordings to validate the results. They concluded that CNN is most suitable when using raw EEG waveforms images for automated seizure detection. In 2019 IBM come forward with SeizureNet [17]. It was able to classify seizures at 98.4% accuracy. SeizureNet's architecture is divided into two sub-networks i.e., a deep convolutional network and a classification network. The dataset used in this paper is TUH EEG Seizure Corpus (v1.4.0).

3. Proposed model

Through this study, we are mainly focusing on performing a comparative analysis of different machine learning algorithms including deep learning techniques. To satisfy our requirement, we have created 8 models using different techniques and trained & tested them on two different EEG datasets- the Bonn dataset [11], and the CHB-MIT dataset [12]. EEG signals are highly unpredictable data with a massive amount of variability on it. It shows different potential at different frequencies, so the best way to analyze an EEG signal is to convert it from the time-amplitude domain into the frequency-time domain. For conversion and feature extraction we have used Wavelet transform (WT). Wavelet transforms outputs both time and frequency of a signal as a result we can use these extracted signals for studying underlying hidden correlations between features. All these steps are executed with the help of mne package available in python. mne is a python tool kit that is exclusively dedicated to medical data analysis. It supports functions like wavelet transforms, down sampling, filtering, etc. To start with, we have used functions like mne. preprocessing and mne. filter for preprocessing our EEG signals. With the help of mne.time-frequency sub-module we have created a time-frequency representation of data. Later with the mne features. feature extraction library, we have extracted prominent features from the data.

Our proposed method includes usage of 8 models, where four of them make use of classical machine learning algorithms like Logistic regression, Naive Bayes, Random Forest, KNN, while the remaining four uses deep learning techniques like an artificial neural network, Convolutional neural network, Long short term memory (LSTM) and Autoencoder. An overview of our study is in figure 1 shown below.

3.1. Classical machine learning models

We have used four supervised machine learning algorithms for the classification task.

3.1.1. Logistic Regression.
Logistic regression classifies the data by learning weights(w) and bias term(b) from the training data. Each input feature (x) is associated with a weight, which is a real number. The weight defines how important an input feature is for the classification decision. After learning the weights and bias from the training data we make use of equation (3.1) to decide on test data.

$$y = \sigma(z) = \frac{1}{1+e^{-z}} = \frac{1}{1+\exp(-z)} \tag{3.1}$$

Where z expresses the weighted sum of the evidence for the class as shown in equation (3.2).
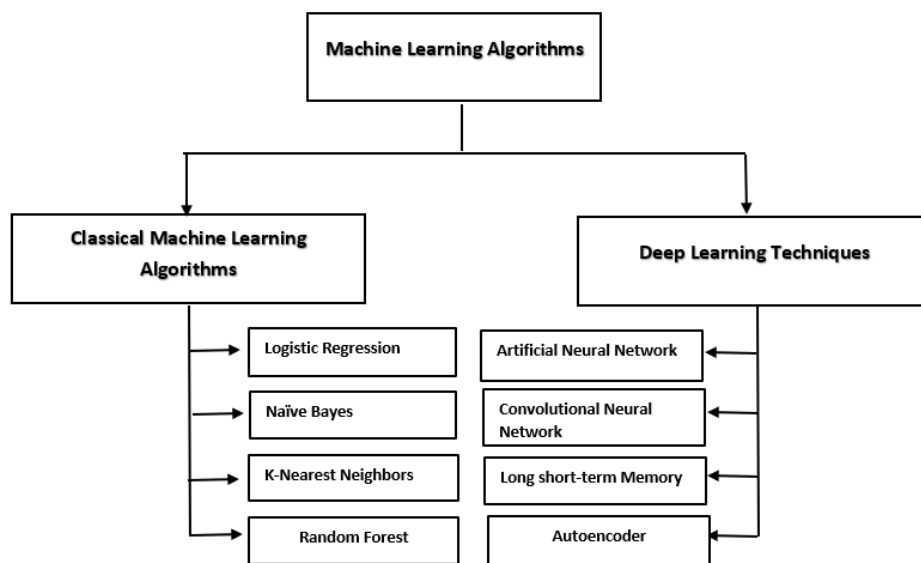
$$z = \omega \cdot x + b \tag{3.2}$$

Figure 1. Overview.

3.1.2. Naive Bayes. The naive Bayes classifier is based on Bayes' theorem. It generates a probabilistic model. It is called naive as it assumes that all the features are independent of each other. If we assume the likelihood of the features to be Gaussian, then the conditional probability would be expressed as stated in equation (3.3).

$$P(x_i|y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left(-\frac{(x_i - u_y)^2}{2\sigma_y^2}\right) \tag{3.3}$$

3.1.3. Random Forest. A random forest is an ensemble of decision trees, which are machine learning algorithms able to perform both classification and regression. This is a part of decisionalgorithms which make use of tree representations.

3.1.4. K Nearest Neighbors. This supervised learning algorithm creates clusters based on the similarity between points and classifies the new points into the class of the points, that is most

similar to it.

3.2. Deep learning models

Deep learning techniques, make use of neural network algorithms that work on the principles of forwarding propagation and backpropagation [13]. Initially, information is passed into the input layer, it is then transferred into the hidden layers. Interconnections between input layers and hidden layers assign a weight to each input. This weight is multiplied with every input. After that, a bias term is added to this. This is then passed to an activation function. Activation determines the limit of the output range, and it decides which features should be selected for appropriate output. Outputs from hidden layers are aggregated in the output layer and it is passed to an activation function to produce output. Later this output is tuned with the help of backpropagation. Neural networks can be again divided into multiple classes based on the methodology they adopt for solving applications.

3.2.1. Artificial Neural Network.
One of the models we had used for our comparative study is a basic artificial neural network that makes use of 12 fully connected layers. This model is a multi-class classification model that works on ReLu and SoftMax activation functions. This model is tested on two datasets to study the performance of the model on different amounts of data.

3.2.2. Convolutional Neural Network.
Convolutional Neural Network is a kind of artificial neural network that works on principles of forward propagation and backward propagation [14]. Our proposed model makes use of a 1 D convolutional network. It makes use of ReLu and SoftMax activations. To avoid overfitting, we have incorporated a drop out of 0.2. To increase computation speed, we have added max-pooling layers in the model. For the same purpose, we have replaced the fully connected layer in the output with global average pooling.

3.2.3. Long term short Memory (LSTM).
Long-term short memory (LSTM) is an upgraded version of RNN with memory elements. RNN suffers with long term dependency problem, this is overcoming in LSTM with the help of cell state. As a result, LSTM is suitable for time series analysis of larger datasets. Our network architecture makes use of two LSTM layers and two fully connected layers. A dropout of 0.3 is induced in the network to avoid overfitting. We have used tanh, sigmoid, and SoftMax activations in our model.

3.2.4. Autoencoder.
Autoencoder is an unsupervised machine learning algorithm that works based on an encoder-decoder circuit [15]. Autoencoder is commonly used for representation learning and it works on the principle of dimensionality reduction. We have implemented a denoising autoencoder by adding noise into our original data, then we have induced a gaussian dropout of 0.3 in our model to avoid overfitting and to induce faster computation. The proposed model makes use of 6 fully connected layers that uses tanh as the activation. Mean Square Error is used as a metric for calculating the efficiency of the model.

4. Dataset

4.1. Bonn Dataset (primary dataset)

Bonn university dataset [11] contains EEG signals sampled at 173.61 Hz. It is recorded using a 128-channel amplifier system It is mainly used for seizure event detection. EEG recordings were done using standard electrode placement which contains 10-20 electrodes. This dataset contains five sets of data from 100 channels. This is named from A TO E.

4.2. CHB-MIT

The CHB-MIT dataset is a publicly available epilepsy data set. It contains 686 scalp encephalogram recordings of 23 patients treated at the Boston Children's Hospital with intractable seizures. The dataset contains 2 recordings of each patient 1.5 years apart. This data set contains 969 hours of EEG recordings with 173 seizures. Subjects include males, females, and children aged between 10-22.

5. Experiment Results

We have created 8 machine learning models which include 4 classical machine learning techniques and 4 deep learning models. We have tested all these models on 2 different datasets which include the Bonn dataset and CHB-MIT for early prediction of epileptic seizures. To find the best model, we have performed 3 types of analysis. In the first phase, we had chosen the Bonn dataset as our primary dataset and applied all 8 algorithms on it to find the best performing classical machine learning algorithm and deep learning algorithm, later to validate the results we have applied the same models on 1 more dataset and made a comparative study to analyze which algorithm is performing better among all models. In the third phase, we had made a comparative analysis of the best performing classical machine learning and deep learning models. To frame our deep learning models, we have used Keras from the tensor flow. It incorporates data generators as root class. For data preprocessing and data extraction, we have used mne package. It has specialized modules for analyzing medical data. Another package we have used is scikit learn. It offers many functionalities like data splitting, in-built modules for computation, etc. Then we had made use of basic python packages like NumPy, pandas, matplot library and so on for analyzing our data sets. We had used precision, recall, sensitivity and specificity, and accuracy as evaluating metrics for our models.

5.1. Prediction on Bonn dataset

We had considered the Bonn data set as our primary dataset. Our first aim was to find the best classical machine learning algorithm for EEG data. Results obtained for 4 classical models are shown in table 1.

5.1.1. Classical Machine learning algorithms.

Table 1. Results-Classical Machine learning algorithms.

Model	Accuracy	Precision	Recall	Specificity
Logistic R	0.80	0.82	0.40	0.56
K N N	0.93	0.92	0.76	0.79
Naive B	0.80	0.65	0.62	0.61
Random F	0.97	0.94	0.92	0.95

For a better understanding of performance, we have used the ROC curve (figure 2) to analyze our models.
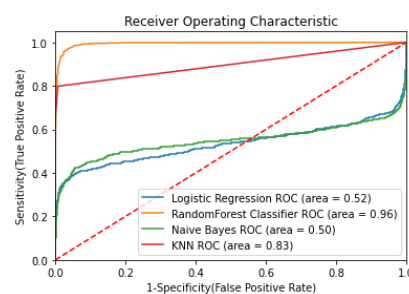

Figure 2. ROC curve.

From this analysis it is understandable that Random Forest Algorithms performs better among all classical algorithms.

5.1.2. Deep Learning Algorithms. We have developed 4 neutral networks for analyzing the Bonn dataset, for a better understanding of model performance, we have created a graph that showcases changes in loss and accuracy with respect to variation of epochs.

Artificial Neural Network The graph in figure 3 shows the variation in accuracy with respect to increase in epochs. From graph, we can understand that accuracy for training data set is increasing constantly, while the accuracy of validation set is initially increasing and later it is remaining constant. This is due to overfitting.
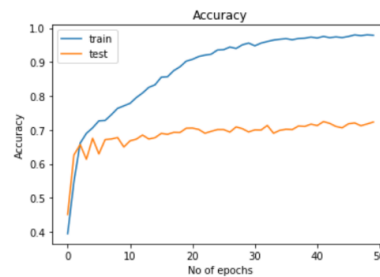
Figure 3. Number of Epochs- Accuracy Curve.

A similar thing can be observed in the Loss-Number of Epochs curve in figure 4, Here loss of training data is decreasing with an increase in the number of epochs. But the loss in validation set initially decreases and later it increases.
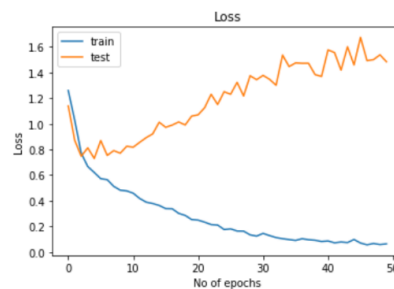

Figure 4. Number of Epochs- Loss Curve.

Convolutional Neural Network figure 5 illustrates the accuracy curve of the Convolutional network. The accuracy of the model increases in both training data and validation data, With the increase in the number of epochs.
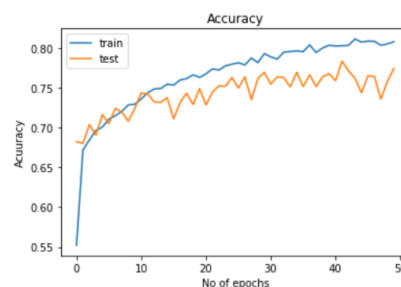

Figure 5. Number of epochs-Accuracy curve.

7

Similarly, loss of both sets are decreasing with an increase in epochs in figure 6. This shows CNN is a good model for the EEG dataset.
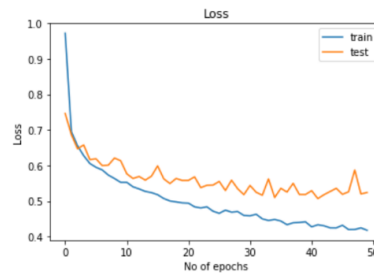

Figure 6. Number of epochs-Loss curve.

Long short term Memory In Accuracy curve (figure 7) of Long Short Term Memory (LSTM) model, we can see a steep increase in accuracy for both training and validation data. This indicates the model is performing well for EEG data.
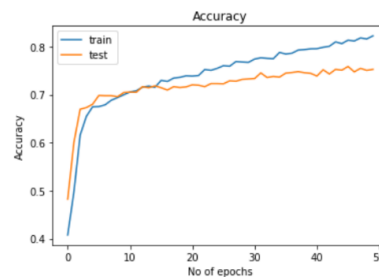

Figure 7. Number of epochs-Accuracy curve.

Loss curve of LSTM model in figure 8 showcase a decrease in loss with respect to an increase in epochs. From that, we can conclude LSTM model is performing well on datasets.
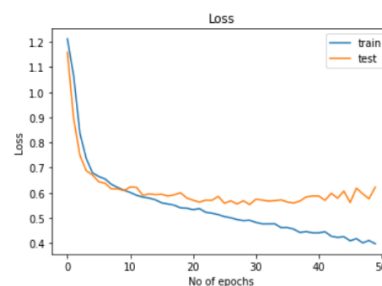

Figure 8. Number of epochs-Loss curve.

Autoencoder Here we have used MSE (mean square error) as metrics to determine the efficiency of the model. The graph in figure 9 indicates MSE error variation with an increase in sample for 2 classes. Here blue bar indicates samples of normal persons and red indicates samples of persons with seizures.
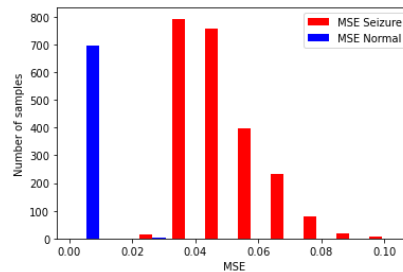
Figure 9. MSE-No of curves.

Comparison of Supervised Neural Networks The graph in figure 10 represents an accuracy comparison of Artificial Neural Network, Convolutional Neural Network, and LSTM. From this graph we can understand LSTM and CNN performed well on the Bonn dataset.
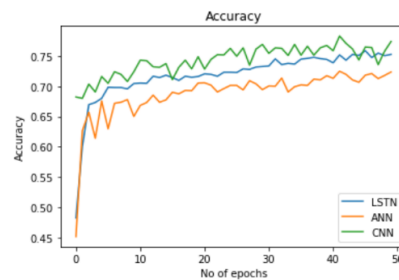

Figure 10. Epoch Accuracy curve of 3 models.

Results obtained from the above 4 deep learning models are presented in table 2. From values we can understand both Convolutional Neural Network and Long Short Term was equally good for analyzing our Bonn EGG data.

Table 2. Results-Deep learning algorithms.

Model	Accuracy	Precision	Recall	Specificity
ANN	0.96	0.97	0.85	0.87
C N N	0.97	0.97	0.95	0.94
LSTM	0.98	0.97	0.97	0.97
Auto Encoder	0.95	0.96	0.88	0.92

Our aim was to find the best machine learning models that perform well on EGG data. To test its efficiency, we have tried same models on 1 more data set. Its result is depicted in table 3.

5.2. Prediction on CHB-MIT dataset

Table 3. Results Machine learning algorithms.

Model	Accuracy	Precision	Recall	Specificity
Logistic R	0.81	0.83	0.6	0.78
K N N	0.91	0.95	0.78	0.89
Naive B	0.82	0.58	0.40	0.6
Random F	0.97	0.94	0.93	0.96
ANN	0.94	0.95	0.83	0.84
C N N	0.96	0.96	0.93	0.92
LSTM	0.98	0.98	0.97	0.97
Auto Encoder	0.97	0.96	0.90	0.92

From these experiments, we can understand Long Short-Term Memory (LSTM) model is the best deep learning model among all deep learning architectures.

5.3. Comparison of Random Forest with LSTM

From the above results can conclude Random Forest is the best classical machine learning to analyze EEG data. Similarly, Long Short-Term Memory (LSTM) is the best deep learning algorithm for EEG datasets. Random Forest is less computationally expensive when compare with the Long Short-Term Memory (LSTM) model. Long Short-Term Memory (LSTM) model performance can be effectively increased with the help of GPU acceleration. Random Forest supports variability of data, while LSTM is more suitable for time-series data. Long Short-Term Memory (LSTM) model has a provision for avoiding long dependencies this makes our predictions more accurate when compared with other models. The main problem with Long Short-Term Memory (LSTM) model is, it requires a large amount of data for training purposes.

6. Conclusion

We have conducted a comparative study of different machine learning techniques to identifythe best machine learning algorithm identifying best performing epileptic seizures prediction method. From our analysis of 4 classical machine algorithms - Logistic Regression, Random Forest, and KNN we concluded that Random Forest is the most cost-effective method to analyze EEG data in case of limited training data. Random Forest supports variability in data. But is Random Forest is not sufficient for analyzing big data, In that case, we have to use neural networks. From our analysis, we had concluded that LSTM is the best deep neural network among Artificial Neural Network, Convolutional Neural Network, and Autoencoders. LSTM has a memory cell, and it avoids long-term dependency problems. LSTM is the best model for seizure prediction of big volume time-series big data. It can achieve high sensitivity and specificity when compared with other models. However, there is still an opportunity to get better in numerous perspectives. In the future, we can incorporate ensemble learning in seizure prediction by combining LSTM with other efficient models. Other than that future research can be done to reduce the number of parameters learned in Deep Learning Models.

Acknowledgments

References

[1] World Health Organization: Who. "Epilepsy". In: *World Health Organization: WHO* (June 2019). URL: https://www.who.int/news-room/fact-sheets/detail/epilepsy.

[2] Tara L. Skarpaas, Beata Jarosiewicz, and Martha J. Morrell. "Brain-responsive neurostimulation for epilepsy (RNS® System)". In: *Epilepsy Research* 153 (2019), pp. 68–70. ISSN: 0920-1211. DOI: https://doi.org/10.1016/j.eplepsyres.2019.02.003. URL: https://www.sciencedirect.com/science/article/pii/S0920121118305758.

[3] V. G. Sujadevi et al. "Anomaly Detection in Phonocardiogram Employing Deep Learning". In: *Computational Intelligence in Data Mining*. Singapore: Springer, July 2018, pp. 525–534. ISBN: 978-981-10-8054-8. DOI: 10.1007/978-981-10-8055-5_47.

[4] Vedha Krishna Yarasuri, Gowtham Kishore Indukuri, and Aswathy K. Nair. "Prediction of Hepatitis Disease Using Machine Learning Technique". In: *2019 Third International conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*. IEEE, Dec. 2019, pp. 265–269. DOI: 10.1109/I-SMAC47947.2019.9032585.

[5] A. Krishna Jothi and Poornima Mohan. "A Comparison between KNN and SVM for Breast Cancer Diagnosis Using GLCM shape and LBP Features". In: *2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT)*. IEEE, Aug. 2020,pp. 1058–1062. DOI: 10.1109/ICSSIT48917.2020.9214235.

[6] Anggiat Saroha Sianipar, Herman Mawengkang, and Suherman. "Performance Analysis of Classifier in Detecting Epileptic Seizure Based on Discrete Wavelet Transform". In: *2020 3rd International Conference on Mechanical, Electronics, Computer, and Industrial Technology (MECnIT)*. 2020, pp. 137–142. DOI: 10.1109/MECnIT48290.2020.9166647.

[7] R. Chinmayi et al. "Extracting the features of emotion from EEG signals and classify using affective computing". In: *2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*. IEEE, Mar. 2017, pp. 2032–2036. DOI: 10.1109/WiSPNET.2017.8300118.

[8] S. S. Poorna et al. "EEG Based Control - A Study Using Wavelet Features". In: *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*. IEEE, Sept. 2018, pp. 550–553. DOI: 10.1109/ICACCI.2018.8554879.

[9] Suvadeep Bose et al. "EEG signal analysis for Seizure detection using Discrete Wavelet Transform and Random Forest". In: *2017 International Conference on Computer and Applications (ICCA)*. 2017, pp. 369–378. DOI: 10.1109/COMAPP.2017.8079760.

[10] D. Selvathi and V. Krishna Meera. "Realization of epileptic seizure detection in EEG signal using wavelet transform and SVM classifier". In: *2017 International Conference on Signal Processing and Communication (ICSPC)*. 2017, pp. 18–22. DOI: 10.1109/CSPC.2017.8305848.

[11] R. G. Andrzejak et al. "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: dependence on recording region and brain state". In: *Phys Rev E Stat Nonlin Soft Matter Phys* 64.6 Pt 1 (Dec. 2001), p. 061907.

[12] Ali Hossam Shoeb. "Application of machine learning to epileptic seizure onset detection and treatment". PhD thesis. Massachusetts Institute of Technology, 2009.

[13] Abdel-Nasser Sharkawy. "Principle of Neural Network and Its Main Types: Review". In: *Journal of Advances in Applied Computational Mathematics* 7 (Aug. 2020), pp. 8–19. DOI: 10.15377/2409-5761.2020.07.2.

[14] Ahmed Shamsaldin et al. "The Study of The Convolutional Neural Networks Applications". In: *UKH Journal of Science and Engineering* 3 (Dec. 2019), pp. 31–40. DOI: 10.25079/ukhjse.v3n2y2019.pp31-40.

[15] Richard Preen, Stewart Wilson, and Larry Bull. "Autoencoding with a Classifier System". In: *IEEE Transactions on Evolutionary Computation* PP (May 2021), pp. 1–1. DOI: 10.1109/TEVC.2021.3079320.