Comparative analysis of convolutional neural networks and rule-based techniques for epileptic seizure detection from electroencephalograph signals using a text classification approach*

Temitayo Balogun[1,†], Raphael Akinyede[1,†] and Samuel Faluyi[2,†]

[1] *Federal University of Technology, Akure, Nigeria*
[2] *Ekiti State Polytechnic, Isan-Ekiti, Nigeria*

Abstract
EEG (Electroencephalograph) signals can be used to determine whether a person is going to have a seizure or not. EEG has proven essential in the early detection of epileptic seizures. To detect epileptic seizures using EEG signals, several machine learning models have been developed. However, others claim that the traditional rule-based approach is just as effective. This study aims to disprove this claim and compare the performance of a rule-based technique and a machine learning approach. Because of how closely it resembles the human brain, the neural network was chosen as the machine learning approach. The dataset was obtained from the open source, freely used Temple University Hospital Abnormal (TUAB) EEG Corpus. The rule-based technique had an accuracy of 85.16% whereas the neural network technique had an accuracy of 98.91% after the data had been taught and tested using both approaches.

Keywords
Neural Network, Rule-Based, Electroencephalograph, Machine Learning, Epileptic Seizures

1. Introduction

Electroencephalography (EEG) has emerged as one of the key methods for studying brainwave patterns [1]. Since the turn of the 20th century, research on using EEG signals to diagnose neurological illnesses has been ongoing and is still going strong today. The scope of EEG research has greatly increased [2]. The range of this research area has expanded to the point that several connections between motor activity, mental state, and mental activity have been made. Despite these advancements, there are still a few data points that can be derived from the EEG. Epileptic seizures and occasional cerebrum movement can be detected by EEG [3,4]. Thus, playing a crucial role in understanding the correlation of both epilepsy and brain damage.

The EEG includes a few characteristics, including high-dimensional spatial and temporal components that might not be prepared by conventional regular measurement techniques. Since high-dimensional EEG data can aggregate into patterns for classification, efficient detection hence requires the use of high-quality machine learning models [1]. The importance of using computer-aided devices and cutting-edge internet of medical things (IoMT) technologies to identify and categorize atypical brain processes and seizures for efficient observation, inspection, analysis, and diagnosis cannot be overstated [4].

Over the years, neuroscience has attempted to employ medical data to manually identify seizures early, but their efforts have been unsuccessful. Information technology, or IT, has aided in the development of models that might use patient data from the past to swiftly identify these epileptic

IDDM'24: 7th International Conference on Informatics & Data-Driven Medicine, November 14 - 16, 2024, Birmingham, UK
* Corresponding author.
† These authors contributed equally.
tebalogun@futa.edu.ng (T. Balogun); roakinyede@futa.edu.ng (R. Akinyede); sgfaluyi.edu.ng (S. Faluyi)
0000-0003-0622-8965 (T. Balogun); 0000-0002-5544-8529 (R. Akinyede); 0000-0002-2662-1599 (S. Faluyi)

episodes and determine the patient's stage. One of the numerous benefits of this early diagnosis is that patients can be spared the negative effects of epileptic seizures if they can be foreseen. The goal of this research work is to use a rule-based and deep learning model to predict epileptic seizures using EEG and compare the findings to determine the most effective.

Early seizure detection has significant implications for various medical specialties, particularly neurology. The ability to predict seizures can improve patient safety and quality of life. Given the seriousness of epilepsy, incorporating machine learning models can be a valuable tool for healthcare professionals [5]. Machine learning offers a promising new approach for predicting epileptic seizures. This goes beyond traditional EEG analysis used for seizure detection and classification during medical examinations. By analyzing EEG data, machine learning models can potentially anticipate seizures before they occur.

2. Literature Review

This section discusses some of the projects that were investigated about the use of various machine learning models to detect and categorize epileptic episodes using EEG, along with their objectives and difficulties.

Almustafa [6] classified an epileptic seizure dataset using a variety of machine-learning techniques. Several classifiers were used to categorize the Epileptic Seizure dataset. In their system, the Random Forest (RF) classifier outperformed the Naïve Bayes model, K-Nearest Neighbor, Logistic Regression, J48, Decision Tree, Stochastic Gradient Descent (SGD) and Random Tree classifiers with 97.08% accuracy, ROC = 0.996, and RMSE = 0.1527. Several of these classifiers underwent sensitivity analysis to determine how well they classified the Epileptic Seizure dataset when some of their parameters were altered. The results imply that the accuracy of the classifier can be enhanced by changing a few classifier parameters. For instance, altering the training/testing split increases the random forest classifier's accuracy to 97.35%, altering the SGD classifier's learning rate to 0.1 raises it to 81.97%, and altering the regularization parameter to 10,000 raises it to 81.92%. Additionally, employing only 148 of the 178 features that can be utilized to predict epileptic seizures, good classification accuracy was achieved using the Naïve Bayes classifier feature extraction method which was dependent on the variance of the available features in the epileptic seizure dataset. Additionally, the dataset was predicted using the feature selection attribute variance basis. However, a task limitation was discovered during implementation, which involved working with a massive dataset with a significant number of 178 features. Feature reduction may have been employed with some chosen features to obtain an accurate forecast of elliptic seizure cases.

Lasefr et al. [7] created An Efficient Automated Technique for Epilepsy Seizure Detection Using EEG Signals. In addition to analyzing the characteristics of brain signals at different phases, the study developed a method for identifying epileptic signals. They used signal processing methods to identify epilepsy in the EEG signal. To make sure that the operational frequency of the signal matched the oversampling requirements, the signal processing procedure started at a sampling rate of 178.6 Hz. The frequency spectrum is then compressed to less than 200 Hz by dividing the signal into five distinct signal levels, each employing a different wavelet filter. There is still reliance on time domain and frequency domain features because they were used to extract properties from an EEG signal rather than statistical data. These characteristics are found in the EEG data utilizing K-Nearest Neighbor (KNN), Support Vector Machine (SVM), and Artificial Neural Networks (ANN) to diagnose epilepsy. Different sets of brain signals were tested, and the results showed that the signals behaved normally and epileptically during a seizure. The KNN had the highest accuracy (95.68%), next was the SVM (94.92%), and the ANN (95.03%).

Shoka et al., [8] developed an automated seizure diagnosis system based on EEG data feature extraction and channel selection. There were five steps in the proposed approach. The first step was to use the variance parameter to choose the most affected channels to reduce dimensionality. The second phase was feature extraction, which involved extracting the 11 most significant features from the selected channels. The 11 features collected from each channel were then averaged in the third phase. The average characteristics were then classified using the classification process in the fourth phase. Finally, the proposed algorithm was cross-validated and tested by separating the dataset into training and testing sets. A comparison of seven classifiers was offered in the research work. Two techniques of testing were used to evaluate these classifiers: random case testing and continuous case testing. In the random case procedure, the KNN classifier outperformed the other classifiers in terms of precision, specificity, and positive prediction. Despite this, the ensemble classifier outperformed the other classifiers in terms of sensitivity and miss rate (2.3%). The ensemble classifier had greater metric parameters than the other classifiers in the continuous case test technique. Furthermore, the ensemble classifier was able to correctly detect all seizure occurrences.

The K Nearest Neighbors (KNN), Naive Bayes, J48, Logistic Regression, and Random Forest approaches were employed in producing predictions and they were analyzed together with Random Forest showing the highest accuracy of roughly 97.08%, according to AlMustafa's [6] research. The KNN model has the highest accuracy of 95.68% in a comparison by Lasefr et al. utilizing the K Nearest Neighbour, Support Vector Machine, and Artificial Neural Network [7]. To determine which method predicted better, Shoka et al. [8] evaluated SVM, Logistic Regression, Decision Tree, Ensembled Model, and KNN; the decision tree and Ensembled Model had a joint maximum accuracy of 90%. But the goal of this study is to expand on the work of some of these researchers and take things even further. This research will examine whether a machine learning model will perform better than an IF… THEN principle because no scholars have compared machine learning with a rule-based system. This study also intends to use an artificial neural network, which was one of the machine learning models used by [7]. This model was chosen because it somewhat resembles the human brain and is thought to be as intelligent as other models [9]. In comparison to [7] and [8], our present research endeavor likewise seeks to achieve higher accuracy.

3. Methodology and system analysis

This section gives a thorough review of an epileptic seizure detection system. Figure 1 shows the steps of a typical system. EEG data collection, preprocessing, feature extraction, classification, and performance analysis and evaluation are the steps carried out in this research work.

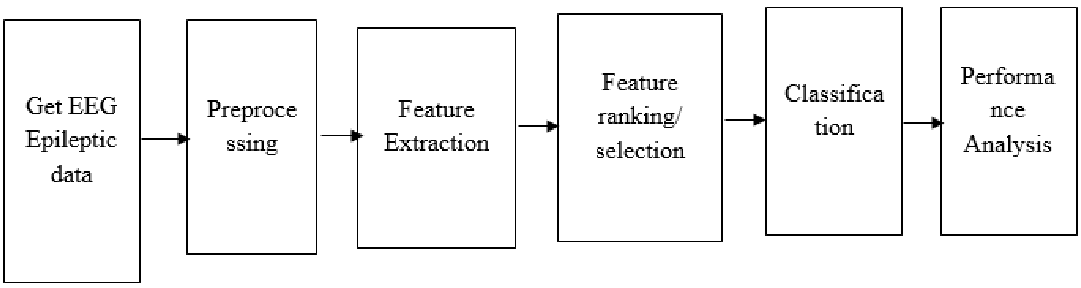

Figure 1: Approach to building the epileptic seizures detection system

3.1. Dataset Collection

For this research, both scalp EEG recordings (EEG) and intracranial EEG recordings (iEEG) were employed by using the 10-20 system, which places electrodes on the surface of the head at equal distances. This method is frequently employed for scalp EEG recordings [10,11]. Intracranial electrodes are placed inside the skull to locate the epileptogenicity zone in the brain when clinical, structural, and functional data are acquired before implantation [12]. Prior investigations made use of the information and data collected from epilepsy patients and analyzed before epileptic procedures to build local databases. The importance of these factors was constrained, which hampered the specificity evaluation in interictal signals. These factors included small sample sizes, short time intervals preceding seizures, and modest seizure movements. As a result, to accurately and effectively assess the sensitivity and specificity of algorithms, long-term signals from several seizures must be recorded [13]. Numerous epilepsy research projects have recently used the Andrzejak database [14] from the Department of Epileptology at the University of Bonn in Germany and the Freiburg database from the Epilepsy Center of the University Hospital of Freiburg in Germany (The University of Freiburg, EEG Database at the Epilepsy Center of the University Hospital of Freiburg in Germany) [15]. The Temple University Hospital Abnormal (TUAB) EEG Corpus, which is open source and freely used, was employed as the dataset for this study. The TUAB dataset is a publicly available data that was downloaded. The dataset was a compressed tar archive named TUAB_txt_relabelled.tar, which contained text documents organized into directories representing different classes. The data contained text in different files which discussed the clinical history of the patient, their medications, an introduction to previous procedures of the patient, a description of their record, their HR, clinical correlation and target values. The dataset was extracted using standard file extraction techniques. The dataset contained 2992 data with 1515 normal and 1477 abnormal cases. However, 2716 records of the data were used for training the models with 1365 normal and 1,351 abnormal cases making up the training dataset as shown in figure 2, while 276 data was used for evaluation with 150 normal and 126 abnormal instances made up the evaluation dataset. Patients who are neither epileptic nor suffering from a brain disorder are represented by Normal, whereas those with epileptic seizures or brain disorders were represented by Abnormal.

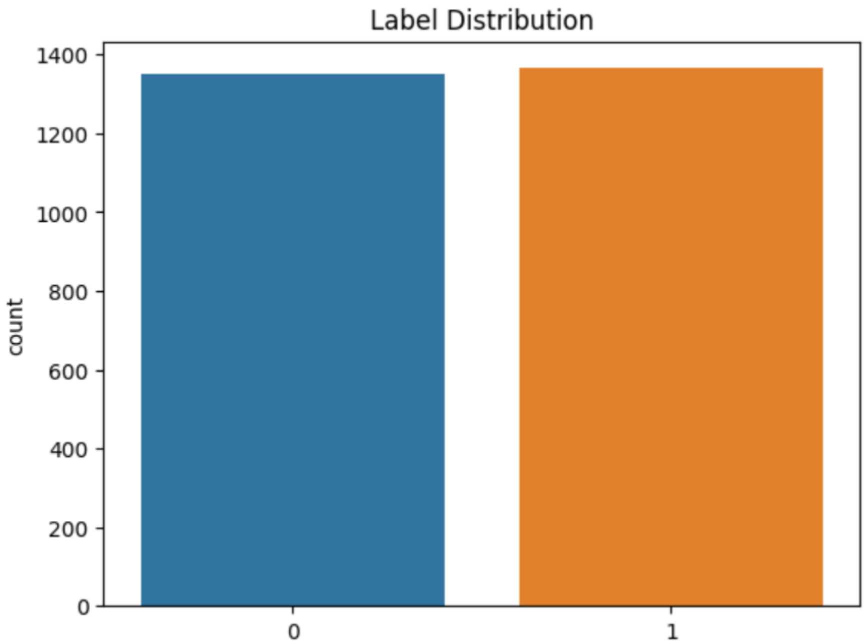

Figure 2: Distribution of training data

3.2. Data preprocessing and feature extraction

Biomedical signals are commonly contaminated with different kinds of noise and artifacts during data collection and processing, which has a massive effect on feature extraction quality [16,17,18]. The essence of denoising and preprocessing cannot be overemphasized with different methods and algorithms developed over time to remove artifacts and noise, making the data more reliable for subsequent processing and analysis [19].

To extract relevant features, Text vectorization was performed using the TextVectorization layer in TensorFlow, which converts raw text into numerical representations suitable for model input. Two vectorization layers are employed which are Binary and Integer Vectorization. The binary vectorization converts text into binary vectors indicating the presence or absence of vocabulary terms while integer vectorization converts text into sequences of integers where each integer corresponds to a vocabulary term. The maximum vocabulary size was set to 10,000 terms, and sequences are padded to a maximum length of 250. The text was further cleaned by transcoding the text into a standard format to eliminate non-standard characters.

3.3. Classification techniques

Accurately identifying different seizure types depends on the quality of features extracted for classification. These features serve as a guide, enabling the classifier to differentiate between various seizure types and normal brain activity. Classifiers, which are decision-making algorithms, analyze these features to establish boundaries between different seizure categories.

The classification process typically involves two stages: training and testing. During the training phase, a selected classification method—ranging from basic thresholding to advanced machine learning algorithms—learns from a labeled dataset that includes extracted features and their corresponding seizure types. Once trained, the classifier can categorize new, unseen data based on the patterns it has learned.

Various techniques can be used for seizure classification, including statistical analyses like clustering, machine learning algorithms, and, more recently, deep neural networks. This study will focus on two specific methods: rule-based classification and convolutional neural networks.

3.4. Rule-based technique

The rule-based strategy is an expansion of Boolean logic. It excels at giving exact responses to problems involving the manipulation of numerous variables. It is utilized in this research to offer a more specialized detection. The dataset employed in this work is made up of a sequence of decision-supporting IF-THEN statements, and it acts as the database engine whereby the approach generates predictions. The method applies pre-defined techniques to the values it receives from the dataset as input. The outputs from the dataset are then loaded into a pre-programmed procedure to produce the prediction.

3.5. Convolutional neural network

This study employed the Convolutional Neural Network model. The Convolutional Neural Network (CNN) was developed with inspiration from the biological concept of a neural network [26]. The model consists of five layers: an embedding layer which transforms integer sequences into dense vectors of size 128, a Conv1D layer which applies convolutional operations to extract features from the text, a MaxPooling1D layer which reduces dimensionality by downsampling the feature maps, a flatten layer which converts the 2D feature maps into 1D and a dense layer which does the classification using the fully connected layers with ReLU activation and a final softmax layer. When employing neural networks, normalized data is necessary for a higher-performing model. After the data was standardized, the proposed model was created using the neural network technique. 20% of the dataset was used for testing, and 80% of the dataset was used for training. The model was compiled with the Adam optimizer and a sparse categorical cross-entropy loss function.

4. Experimental setup and discussion

It was necessary to collect data, train the data obtained, process the data, and develop the systems to build the epileptic seizures detection system utilizing EEG. To confirm system performance and assess how helpful and accurate the detection systems were, thorough analysis was carried out. The TUAB (Temple University Hospital Abnormal) EEG Corpus, whose dataset is open source and freely used by the public was employed for the study. A total of 2716 data was used for training the models with 1365 normal and 1,351 abnormal cases making up the training dataset, whereas 276 data was used for evaluation with 150 normal and 130 abnormal instances made up the evaluation dataset. Patients who do not have epilepsy or a brain disorder are referred to as normal, whereas those who do are referred to as abnormal.

A portion of the data was utilized for validation in addition to training the system, and the remainder was used for testing. The training process resulted in some losses for the system as well. For maximum effectiveness, the dataset was trained and retrained over 10 epochs for the model to fully understand the data.

4.1. Overfitting

To assess the accuracy of the model, its performance was compared on both training and validation datasets, which helps identify whether the machine learning system is experiencing overfitting or underfitting. Unlike previous studies that typically used two data groups, this dataset for this research was divided into three to provide a more robust check against these issues. Specifically, the training data was split into two parts with 80% for training the model and 20% for validation, while an additional evaluation set is reserved for testing the model's performance after training.

When evaluating the performance of a model, underfitting occurs when the validation accuracy is significantly higher than the training accuracy, indicating that the model is too simple to capture the underlying patterns in the training data. In contrast, overfitting happens when the training accuracy is much higher than the validation accuracy, suggesting that the model has learned the training data too well, including its noise and outliers, and is therefore not generalizing effectively to new data. This three-group approach allows for a more thorough evaluation of the ability of the model to generalize the unseen data.

```
train_and_val_ds = preprocessing.text_dataset_from_directory(dataset_dir/'train', batch_size=32)
raw_test_ds = preprocessing.text_dataset_from_directory(dataset_dir/'eval', batch_size=32)
```

```
Found 2716 files belonging to 2 classes.
Found 276 files belonging to 2 classes.
```

When running a machine learning experiment, it is a best practice to divide your dataset into three splits: train, validation, and test. There are no strict rules, but usually it's best to put most of your data in the training (so that there's plenty to learn from. Let's split the training-and-validation data into 80% training and 20% validation.

```
# Set the size of each subset of data:
n = len(list(train_and_val_ds)) # Number of batches in original 'train' dataset
n_train = int(0.8*n)    # Use about 80% as training data ...
n_val = int(0.2*n)      # and 20% as validation data.
```

Now we're ready to actually make the split.

```
# Split the training data into training, validation:
raw_train_ds = train_and_val_ds.take(n_train)
raw_val_ds = train_and_val_ds.skip(n_train)
```

Figure 3: Data split into three sets; testing, validation and testing.

4.2. System Evaluation

Upon feeding the data into the system for training, the performance of the model is determined to know how much the models have learned, the validation and test data is used to evaluate the effectiveness of the machine learning model. Figures 4 to 6 present the confusion matrix and accuracy report for both the convolutional neural network (CNN) and the rule-based model.

Furthermore, the accuracy of the rule-based model was calculated with an accuracy of 85.16% as shown in figure 7.

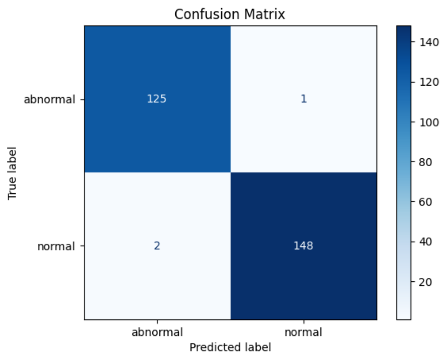

Figure 4: Confusion Matrix for CNN

```
print(classification_report(y_true, y_pred, target_names=train_and_val_ds.class_names))
```

```
              precision    recall  f1-score   support

    abnormal       0.98      0.99      0.99       126
      normal       0.99      0.99      0.99       150

    accuracy                           0.99       276
   macro avg       0.99      0.99      0.99       276
weighted avg       0.99      0.99      0.99       276
```

Figure 5: Classification Report for the CNN model

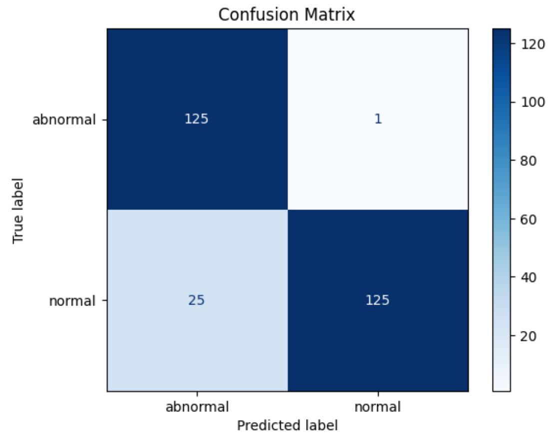

Figure 6: Confusion Matrix for the Rule Based Approach

```
        print("--------------------")
        print(f"The above example was classified with label {predicted_label} but it's actual label is {label_batch[ind].numpy()}.")
        print("--------------------")
        pass

    # Add one to our count of the total number of examples examined.
    n = n+1

print(f"Accuracy = {100*n_correct/n} percent ({n_correct} correct predictions out of {n}).")
```

```
TECHNICAL DIFFICULTIES:  None.
DESCRIPTION OF THE RECORD:  The record reveals a posterior dominant rhythm that at best reaches 9 Hz and is reactive to eye opening.  The patient has a
n excessive amount of beta seen diffusely throughout the recording.  The patient is recorded in wakefulness and sleep, stage 1 and stage 2.  Activating
procedures including hyperventilation and photic stimulation produces no abnormal discharges.
ABNORMAL DISCHARGES:  None.
SEIZURES:  None.
IMPRESSION:  Normal awake and sleep EEG.
CLINICAL CORRELATION:  No seizures or epileptiform discharges were clearly seen.  Please note that the patient's excessive beta may be due to medicatio
n effect usually seen with benzodiazepine or barbiturate use.



--------------------
The above example was classified with label 0 but it's actual label is 1.
--------------------
Accuracy = 85.1620029455081 percent (2313 correct predictions out of 2716).
```

Figure 7: Report of the rule-based approach

Evaluating the accuracy of the testing data in relation to the validation data was a key criterion for assessing overfitting or underfitting. In this case, there is no evidence of either issue, as the accuracy of the validation data closely matches that of the testing data, indicating that the system is performing well.

4.3. Comparative Analysis of the proposed system with existing works

The performance of the system was evaluated against that of other existing systems. This study aimed to build on the work of previous scholars who have conducted similar tasks, thus contributing to the advancement of this research field in practical applications. Consequently, the accuracy of this research was compared to that of three other studies, and it was determined that our work performed well in relation to those that have tackled similar issues.

The evaluation results of the proposed system were compared with previously published works, and these findings are presented in Table 1.

Table 1
Comparative analysis of the accuracy of the system with existing works

Authors	Approach/Technique	Accuracy
AlMustafa, [5]	KNN	0.9523
	Naïve Bayes	0.9573
	J48	0.9482
	Random Forest	0.9708
	Logistic Regression	0.9197
Lasefr et al., [6]	KNN	0.9568
	SVM	0.9492
	ANN	0.9503
Shoka et al., [7]	SVM	0.88
	Logistic Regression	0.84
	Decision Tree	0.90
	Ensembled Model	0.90
	KNN	0.80
Current Research	CNN	0.9891
	Rule Based	0.8516

5. Conclusion

In this study, two methods for predicting epileptic seizures were examined, comparing a rule-based approach with a machine learning approach. While numerous machine learning techniques have been explored in previous research, this study used a neural network due to its effectiveness in modeling brain functions, which is crucial for verifying EEG data. The neural network is widely recognized as one of the most effective methods for developing seizure prediction systems. Despite the promising results of earlier studies, there remains significant room for improvement, which this research aimed to address. By identifying the limitations in previous work, we sought to enhance the predictive accuracy of seizure detection. The result of this study showed that the neural network outperformed the rule-based method significantly as the neural network achieved an accuracy of 98.91% while the rule-based model achieved an accuracy of 85.16%. This difference showcased the superiority of machine learning in making accurate predictions for seizure events. Future studies should consider comparing a broader range of models to identify the most effective methods for predicting epileptic seizures. This research will be invaluable for clinicians and researchers seeking to enhance seizure prediction systems and ultimately improve patient outcomes.

Declaration on Generative AI

During the preparation of this work, the author(s) used Quillbot, Grammarly to: Paraphrase and reword, Grammar and spelling check. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the publication's content.

Acknowledgements

Special thanks to everyone who supported this research in one way or the other.

References

[1] Si, Y., 2020. Machine Learning Applications for Electroencephalograph Signals in Epilepsy: A Quick Review. Acta Epileptologica [online]. 2(5), pp. 1-7. [Accessed 24 April 2024].
[2] Ildar, R., 2020. Review of Neural networks in the EEG signal recognition [online]. Chelyabinsk: South Ural State University, Department of Power Plants Networks and Systems, Russia. [Accessed 24 April 2024].
[3] Rasheed, K. et al., 2021. Machine Learning for Predicting Epileptic Seizures Machine Learning for Predicting Epileptic Seizures. IEEE Reviews in Biomedical Engineering [online]. 14, pp. 139-155. [Accessed 24 April 2024].
[4] Saminu, S. et al., 2021. A Recent Investigation on Detection and Classification of Epileptic Seizure Techniques Using EEG Signal. Brain Sciences [online]. 11(668). [Accessed 24 April 2024].
[5] Faluyi, S., Fapohunda, K., Balogun, T. and Oluwabusayo, O., 2024. Hepatocellular Carcinoma Patients Survival Forecasting Model Using Ensemble Learning Approach. In Data-Centric AI Solutions and Emerging Technologies in the Healthcare Ecosystem [online] (pp. 185-203). CRC Press. [Accessed 8 July 2024]
[6] Almustafa, K. M., 2020. Classification of epileptic seizure dataset using different machine learning algorithms. Informatics in Medicine Unlocked, [online]. 21, pp. 1-7. [Accessed 15 November 2023].
[7] Lasefr, Z., et al., 2017. An Efficient Automated Technique for Epilepsy Seizure Detection Using EEG signals. IEEE [online]. pp. 1-6 [Accessed 25 April 2024].
[8] Shoka, E. et al., 2019. Fast Seizure Detection from EEG Using Machine Learning. 2019 7th International Japan-Africa Conference on Electronics, Communications, and Computations, (JAC-ECC) [online]. 7(1). pp.120-121 [Accessed 24 April 2024].
[9] Balogun, T., et al., 2022. Comparative analysis of deep learning models for the detection and classification of Diabetes Retinopathy. In 2022 5th Information Technology for Education and Development (ITED) [online] (pp. 1-6). IEEE. [Accessed 20 June 2024].
[10] Mike, C. X., 2014 Analyzing Neural Time Series Data: Theory and Practice. [online]The MIT Press: Cambridge, MA, USA. [Accessed 25 April 2024].
[11] Alahi, M.E.E., Liu, Y., Xu, Z., Wang, H., Wu, T. and Mukhopadhyay, S.C., 2021. Recent advancement of electrocorticography (ECoG) electrodes for chronic neural recording/stimulation. Materials Today Communications, 29, p.102853. [Accessed 25 April 2024].
[12] Amer, N.S. and Belhaouari, S.B., 2023. EEG Signal Processing for Medical Diagnosis, Healthcare, and Monitoring: A Comprehensive Review. IEEE Access. [Accessed 25 April 2024].
[13] Ahmed, F. H., et al., 2018. Focal and non-focal epilepsy localization: A review. IEEE Access [online], 6, pp. 49306–49324. [Accessed 21 November 2023].
[14] EEG Time Series, 2014. [online] Available at: http://epileptologiebonn.de/cms/upload/workgroup/ehnertzeegdata.html [Accessed on 24 January 2022].
[15] Andrzejak, R. G., et al., 2012. Nonrandomness, non-linear dependence and non-stationarity of electroencephalographic recordings from epilepsy patients. Phys. Rev. E [online] 86, 046206. [Accessed 25 April 2024. [Accessed 25 April 2024].
[16] Gorjan, D., Gramann, K., De Pauw, K. and Marusic, U., 2022. Removal of movement-induced EEG artifacts: current state of the art and guidelines. Journal of neural engineering, [online] 19(1), p.011004. [Accessed 25 April 2024].
[17] Sweeney, K.T., et al., 2012. Artifact removal in physiological signals Practices and possibilities. IEEE Trans. Inf. Technol. Biomed. [online] 16, pp. 488–500. [CrossRef]
[18] Belkhiria, C., Boudir, A., Hurter, C. and Peysakhovich, V., 2022. EOG-based human–computer interface: 2000–2020 review. Sensors, [online] 22(13), p.4914. [Accessed 25 April 2024]
[19] Mahamune, R. S., and Laskar, S. H., (2019) A review on artefacts removal techniques for Electroencephalogram signals. In Proceedings of the 2nd International Conference on

innovations in Electronics, Signal Processing, Communication (IESC), [online] Shillong, India, pp. 50–53. [Accessed 25 April 2024].
[20] Tariq, H. I., et al., 2019. Loan Default Prediction Model Using Sample, Explore, Modify, Model, and Assess (SEMMA). Journal of Computational and Theoretical Nanoscience, [online]. 16(8), pp. 3489–3503. [Accessed 25 April 2024].