Nigerian Journal of Engineering Science Research (NIJESR).
Vol. 4, Issue 3, pp. 21-37, September, 2021

ISSN: 2636-7114
Journal Homepage: https://www.iuokada.edu.ng/journals/nijesr/

Epilepsy Seizure Classification Using Artificial Neural Network and Linear Discriminant Analysis Algorithm

[1*]Umar. B. U, [2]Muazu M. B, [3]Agajo. J, [3]Kolo. J. G

[1]Department of Computer Engineering, School of Electrical Engineering and Technology, Federal University of Technology, Minna, Nigeria,
buhariumar@futminna.edu.ng
[2]Department of Computer Engineering, Ahmadu Bello University, Zaria
[3]Department of Computer Engineering, School of Electrical Engineering and Technology, Federal University of Technology, Minna, Nigeria
james.agajo@futminna.edu.ng
jgkolo@futminna.edu.ng

**Corresponding author name: Buhari Ugbede Umar (buhariumar@futminna.edu.ng)*

Manuscript History
Received: 10/08/2021
Revised: 15/09/2021
Accepted: 28/09/2021
Published: 30/09/2021

Abstract: *Epilepsy is a condition that affects 8 out of every 1000 persons on the planet. It's a condition that alters the brain's nerve cell activity, resulting in seizures in the patient. For more than three decades, people have been fascinated by the idea of detecting and forecasting epileptic episodes. According to studies, irregular brain activity occurs a few minutes before the onset of a seizure, which is referred to as the preictal stage. Electroencephalography (EEG) is an electrophysiological monitoring technique that records the electrical activity of the brain in order to detect and forecast epileptic episodes. Predicting epileptic seizures before they happen can help prevent them and guarantee adequate seizure control. Many researchers have attempted to anticipate the preictal stage of a seizure, but successful prediction with high sensitivity and specificity remains a difficulty. This research presents a machine learning model for classifying EEG signals into seizure and non-seizure data that uses Linear Discriminant Analysis (LDA) for feature extraction and an Artificial Neural Network for classification. An EEG recording of ten patients was used to test the proposed approach. The data was filtered, and features were chosen using Linear Discriminant Analysis. The data was divided into seizure and non-seizure categories using an artificial neural network. With a classification time of 0.013s to 1s, the model had an overall accuracy, sensitivity, precision, specificity, and F1-score of 86 percent, 69 percent, 80 percent, 96.2 percent, and 72 percent, respectively. The contribution of this research is the introduction of LDA for smart and*

Keywords: Artificial Neural Network*, Linear Discriminant Analysis, Epilepsy, Prediction, Seizure and Electroencephalography*

INTRODUCTION

Epilepsy is a neurological illness that affects millions of people throughout the world. Epilepsy affects around 65 million people on the planet (World Health Organization, 2019). A seizure is a condition that arises as a result of aberrant electrical activity in the brain. Seizures cause uncontrollable physical convulsions and, in rare cases, unconsciousness (Kiral-Kornek *et al.*, 2018).

The onset of seizures is classified as focal, generalized, uncertain onset, or unclassifiable. Sixty percent of persons with epilepsy have focal seizures, thirty percent have generalized seizures, and ten percent of people with epilepsy have episodes with no known onset (Falco-Walter *et al.*, 2018). Seizures are triggered by an aberrant discharge of brain neurons that results in momentary malfunction. Epilepsy and seizures can occur at any age but are more common in children and older people (Daoud & Bayoumi, 2019). Epileptic patients cannot engage in everyday activities such as swimming, driving, hiking, cooking, because of the unpredictive nature of their seizures. They live in constant anxiety and depression as a result of the unpredictable nature of their seizures. Epilepsy patients and their families face stigma and discrimination. Many children with epilepsy are unable to attend school, and adults are denied certain rights such as the ability to work, the right to drive, and marriage. People with epilepsy's human rights are routinely infringed throughout the world. The unpredictive nature of a seizure puts the epileptic person in danger. They could be in the middle of physical activity and be suddenly attacked by a seizure (Maria *et al*., 2018). This drastically affects the patients' quality of life as they cannot work or lead normal lives. People with epilepsy (PWE) are three times more likely than non-epileptic people to die in an accident, and they have a 5% probability of attending the emergency room each year as a result of epilepsy-related injuries (Lasefr *et al*., 2017). Although with the help of medication, 80% of epilepsy seizures can be managed and controlled but cannot be cured and the patients live with epilepsy all their lives, a system that predicts the occurrence of a seizure would better improve the patient's quality of life.

In most circumstances, a professional neurologist monitors and inspects an epileptic patient's EEG signals, This, however, takes time and is inconvenient in cases where EEG recordings have been made for several hours or days. For quantitative analysis and interpretation, quick seizure detection and prediction system that can classify huge EEG data into epileptic and non-epileptic seizures are required (Sharmila & Mahalakshmi, 2017). Various approaches have been developed over the years to automatically classify EEG data into epileptic and non-epileptic seizures without requiring hours of visual review. This is achieved using pattern recognition (Lasefr *et al*., 2017). The steps of feature extraction, feature reduction, and classification are the three steps in pattern recognition. The technique of extracting features is known as feature extraction, it is used for obtaining the signal's hidden characteristics, such as mean, frequency, and amplitude, to enable subsequent signal processing. Feature reduction or dimensionality reduction techniques such as Linear Discriminant Analysis (LDA) and Principal Component Analysis are used to minimize the number of data features (PCA). The data gotten from the dimensionality reduction is used as input for the classifier (Murugavel, 2018). Statistical, frequency domain, and nonlinear factors have been used to detect, predict, and/or detect and predict epileptic seizures using EEG signals utilizing a variety of machine learning and deep learning methods. The EEG signal is obtained by putting small metal discs (electrodes) on the patient's scalp. Electrical impulses are used by brain cells to communicate and they are active all the time, even while sleeping. Epileptic seizures can be detected and predicted by measuring these signals. In the traditional scenario, the selection of features and classifiers is done by trial and error using machine learning techniques (J. Liu & Woodson, 2019). To construct a reliable model, you'll need a solid understanding of signal processing and data mining. With limited data, these machine learning models perform well. As a result of the increased availability of enormous data, deep learning techniques are used, as machine learning techniques may not perform effectively. The majority of simulations in classical machine learning techniques were run in the MATLAB software environment, whereas deep learning models are typically built using the Python programming language and its various open-source tools. This makes it easier for scholars to contribute to different projects. Also, the accessibility of cloud computing has made it easier for researchers to work with big data.

The application of linear discriminant analysis for extracting the relevant features and classification, along with an artificial neural network, would result in an excellent epilepsy prediction system in this research. EEG is a technique for analyzing and studying the electrical activities of the brain. Around 76 per cent of epilepsy sufferers in developing nations cannot afford medical treatment or diagnosis (Varnado & Price, 2020). As a result of this shortcoming, various automated seizure detection and prediction systems have been developed. Feature extraction identifies EEG signal features that best explain seizures and periods preceding seizures. For feature extraction, several prediction systems use time-domain, frequency-domain, or time-frequency domain approaches (Amin *et al.*, 2019). These strategies each have their own set of benefits and drawbacks. This domain features techniques have certain drawbacks since EEG signals are highly variable, and these techniques lack resilience and are easily prone to failure when seizure patterns vary slightly (Acharya *et al.*, 2018). Automated techniques can't handle non-stationary signals like EEG seizure signals because they are so dynamic. Environmental noise, skin impedance, eye blinks, muscle movements, and breathing movements are all sources of noise and artefacts in EEG data (Shanir *et al.*, 2017). The tiny size of EEG datasets is another important disadvantage of automated approaches. This has a negative impact on the test data's robustness of performance. Because the majority of publicly available EEG datasets are only for a few patients and for a short length of time, systems trained on them are not suitable for clinical application. To address these issues, this study offers a seizure classification system with a low signal-to-noise ratio that extracts features using Linear Discriminant Analysis (LDA) and an Artificial Neural Network for classifying data into epileptic and non-epileptic seizure categories. This study proposes a seizure prediction approach for separating EEG datasets into seizure and non-seizure data. UCI Machine Learning Repository, Irvine, CA: The University of California, School of Information and Computer Science, provided the EEG dataset utilized in the creation and testing of the proposed system. To reduce noise and other interference, the datasets were pre-processed. The EEG data were decomposed using Linear Discriminant Analysis, and spike-based characteristics were recovered. The model was created using an artificial neural network. To compare the performance of the network with different numbers of layers, three models with 3, 6, and 9 layers were created. In terms of the system's general accuracy, sensitivity, and specificity, the results were compared to a model constructed only with Artificial Neural Networks.

MATERIAL AND METHOD

This section details the study plan used for this research work, as well as the resources required, the data set, collection strategy, and the methods used to arrive at the conclusions. Fig 1 present the system block diagram.

2.1 Implementation Flow Steps

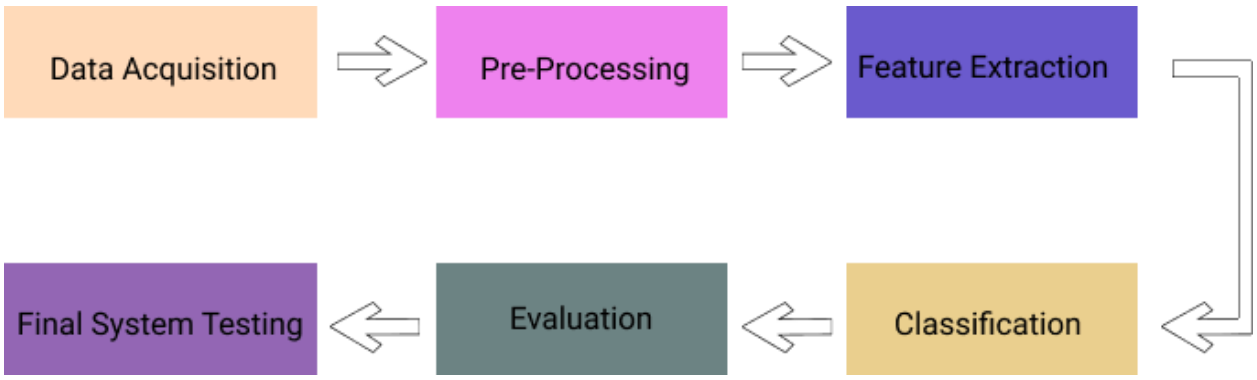

Fig.1 System Block Diagram

2.1.1 Date Acquisition

EEG recordings of ten epileptic patients were obtained from the Neurology & Sleep Centre in Hauz Khas, New Delhi. The data was collected using the Grass Telefactor Comet AS40 Amplification System at a sample rate of 200 Hz. Gold-plated scalp EEG electrodes were put according to the 10-20 electrode placement system during the acquisition. The signals were then split into pre-ictal, interictal, and ictal stages after being filtered between 0.5 and 70 Hz. The collection is divided into five folders, each containing 100 files, each representing a single subject/person. Each file is made up of a 23.6-second recording of brain activity. A total of 4097 data points are sampled from the related time series. Each data point represents the value of an EEG recording at a specific time. There are 500 people in the dataset, each providing 4097 data points lasting 23.5 seconds. Every 4097 data points were separated and jumbled into 23 chunks, each chunk containing 178 data points for 1 second, and each data point represents the value of an EEG recording at a distinct time. The dataset has 23 x 500 = 11500 data points (row), each data point (column) is made up of 178 data points for 1 second, and the last column stands for the label y 1,2,3,4,5. Fig.2 presents the normal EEG signal, while Fig.3 present the Ictal EEG Signal.

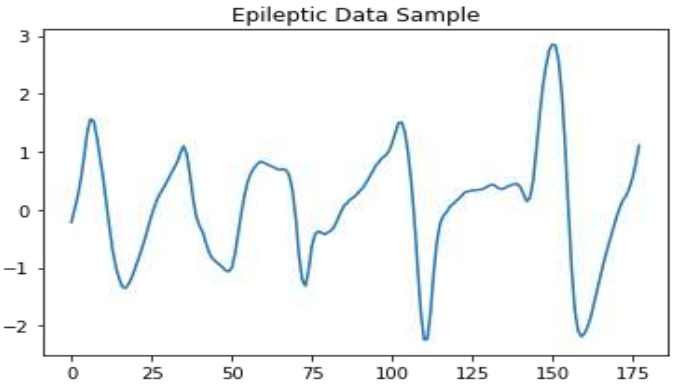

Fig.2 Normal EEG Signal

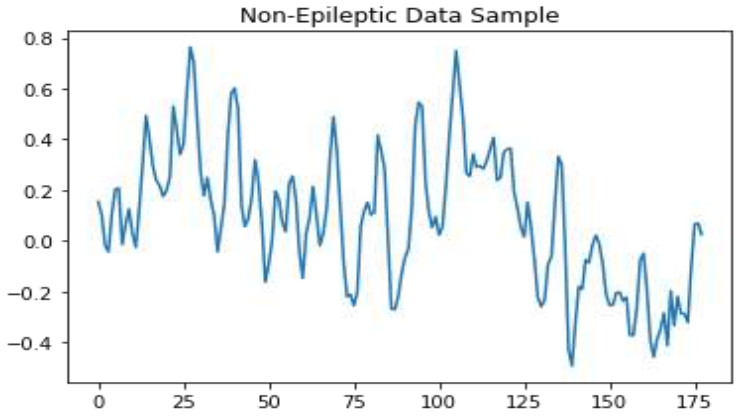

Fig.3 Ictal EEG Signal

2.1.2 Pre-Processing

The term "pre-processing" refers to the process of reducing noise from data to come closer to real brain signals. Before the EEG data was used, it was filtered to remove noise and artificial identification. Many techniques such as band-pass filter, finite impulse response filter, wavelet filter, adaptive filter, etc., are used for filtering. Filtering normalizes the EEG data. Filtering also removes dropouts and corrupted data that occur due to the constraints of implanted electrodes, EEG recording is not possible. It also removes muscle artefacts. Data normalization was performed on the data to convert the dataset to a common scale while preserving variances in value ranges. By

removing the mean and then dividing the result by the standard deviation, normalization was achieved.

$$\text{Nominal Data} = \frac{Data - Mean}{\text{Standard Deviation}} \quad (1)$$

The data dataset features were reduced from 5 to 2. Where '1' stood for 'epileptic' data and '0' stood for 'non-epileptic data.

2.1.3 Feature Extraction

Feature extraction is a procedure that reduces a large set of raw data into smaller groupings before processing. The enormous number of variables in these large data sets necessitates a lot of computational resources to process. Linear Discriminant Analysis (LDA) is a dimensionality reduction technique that is supervised learning and it tries to optimize the distance between each class's meanwhile minimizing the dispersion within each class. As a measure, LDA employs classifications and the distance between them. This is an excellent choice because, when projecting data in a lower-dimensional space, increasing the distance between the means of each class can lead to better classification results. Because LDA assumes that the input data follows a Gaussian distribution (as in this case), using it for data that is not Gaussian may result in poor classification results. Before categorization, the LDA further decreased the features to a more manageable quantity.

2.1.4 ANN Classification

An Artificial Neural Network (ANN) classifier was used for classification. ANNs are computer systems that are based on biological neural networks seen in animal brains. Artificial neurons are a set of connected units or nodes in an ANN that loosely replicate the neurons in a biological brain (Schmidhuber, 2015). The difference between the network's processed output (typically a prediction) and a target output is frequently determined when training a neural network from a given sample. The network then modifies its weighted associations using this error value and a learning strategy. With each adjustment, the neural network will give an output that is quite the target output. After a sufficient number of these modifications, the training can be discontinued based upon specified criteria. Supervised learning is the term for this. The dataset was divided into three groups: training dataset, validation dataset, and test dataset, in the following order: training dataset 70%, while the test and validation dataset were 15% each. The validation and test data were separated from the training data before developing the model to prevent overfitting in the model.

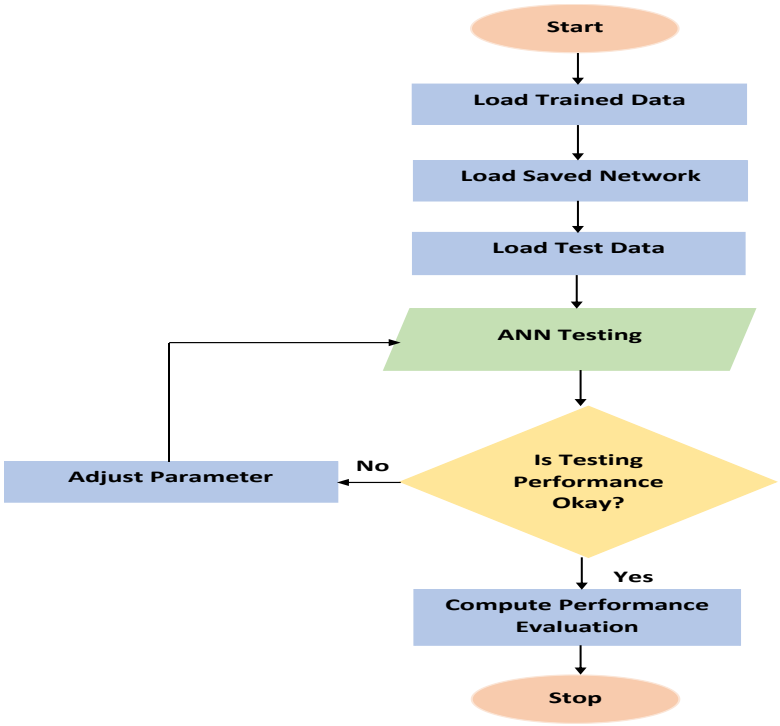

Fig.4 ANN Testing Flowchart

Fig.4 present the trained ANN model that was used for the test data. The performance was then evaluated using sensitivity, specificity and accuracy metrics. Fig.5 present the confusion matrix.

2.1.4.1 ANN Training/ Testing Confusion Matrix

Fig.5 Confusion Matrix

A confusion matrix is used to describe a classification models performance on a set of test data when the true values are knows. The confusion matrix of the model compares the results for true positive, false positive, true negative, and false negative.
Where:

A. True positive (TP): This is the total number of EEG recordings accurately identified as ictal.

B. True negative (TN): This is the total number of EEG recordings accurately identified as seizure-free.

C. False-positive (FP): This is the total number of EEG recordings that have been incorrectly identified as ictal.

D. False-negative (FN): This is the total number of EEG recordings that have been inaccurately predicted as seizure-free.

2.2 Performance Evaluation

To assess a system's performance and quality, several measures must be implemented. As performance parameters for ES predictors offered sensitivity and false prediction rates. The sensitivity of a system is a measure of how well it anticipates or detects seizures. The system's specificity is a measure of how well it classifies normal activities.

2.2.1. Accuracy

Accuracy is the state or degree of precision. Accuracy is the quality of being correct, true, or close to the true value. It is calculated as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (2)$$

2.2.2. Sensitivity or Recall

Sensitivity is the measure of how a system correctly classifies activity. It is calculated as:

$$\text{Sensitivity} = \frac{TP}{TP + FN} \times 100\% \quad (3)$$

2.2.3 Specificity

This is the measure of how a system correctly classifies normal activity. It is calculated as:

$$\text{Specificity} = \frac{TN}{TN + FP} \times 100\% \quad (4)$$

2.2.4 Precision

This is the percentage of relevant instances found among the retrieved ones. It's calculated as follows:

$$\text{Precision} = \frac{TP}{TP + FP} \times 100\% \quad (5)$$

2.2.5 F1 Score

Precision and Recall are weighted in this calculation. It considers both the false positive values and the false negative values. Accuracy is simpler to comprehend, but it is often more beneficial than accuracy when there is an unequal distribution of classes. It is calculated as:

$$\text{F1 Score} = \frac{2 \times \text{Precision} \times \text{Re}\,call}{\text{Precision} + \text{Re}\,call} \quad (6)$$

RESULTS AND DISCUSSION

3.1 LDA Feature Extraction

To better optimize the selected features, LDA was utilized for feature extraction. Three trials were carried out, each with three, six, and nine layers. One experiment was carried out without LDA feature extraction and the result was compared with the optimal result from experiments 1-3.

3.2 Experiment 1

The model developed used LDA for feature extraction and a three-layer Artificial neural network for sorting out the dataset into seizure category and non-seizure category.

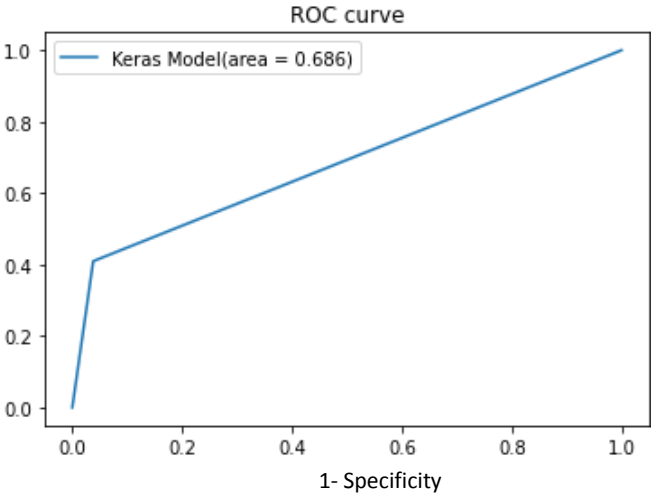

Fig.6 Experiment 1 LDA Convergence Curve

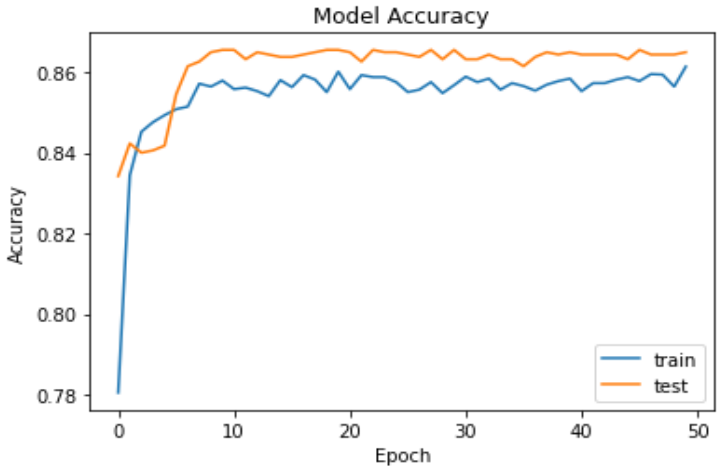

Fig.7 Experiment 1 Model Accuracy

Fig.6, above presents the graph of the sensitivity against the value of 1- specificity or false positive rate of the ANN model. The ROC/ Convergence Curve has an AUC value of 0.686 with an accuracy of 86%. It was discovered that the model's true positive rate increases at a rapid pace in the beginning, when the false positive rate is still close to zero, indicating that the model can capture the critical elements needed to correctly describe and separate seizure and non-seizure data as shown in Fig.7. this shows the accuracy curve against 50 iterations (epochs) of the model for the training and

Umar ***et al.*****, (2021).** ***Epilepsy Seizure Classification Using Artificial Neural Network and Linear Discriminant Analysis Algorithm. Nigeria Journal of Engineering Science Research (NIJESR). 4(3), pp. 21-37***

the test data. The training data has a general accuracy of 87% and the test data has an accuracy of 85%.
That is, the algorithm accurately classifies 86 percent of the data into seizure and non-seizure categories.

3.2.1 ANN Classification Results for Experiment 1

Fig.8 Experiment 1 LDA Confusion Matrix

As shown in Fig.8, 147 is the true positive value, 1313 is the true negative value, 52 is the false positive value, and 212 is the false negative value. It can be shown that LDA correctly classifies 81 percent of non-seizure data while misclassifying 19 percent. Again, it correctly classifies 86.1 percent of seizure data and incorrectly classifies 13.9 percent of seizure data. The classification runtime ranged from 0.013s to 1.s. This is to say that the system showed a fast classification time when the classification iterations were 50.

3.2.2 ANN Validation Results for Experiment 1

Table-1 presents the overall model accuracy, sensitivity, precision, specificity and F1-score for experiment 1. 86%, 69%, 80%, 96.2% and 72%, respectively was achieved. That is, the model properly classifies 96 percent of the data as non-seizure and 41 percent of the data as a seizure.

Table - 1 Summary of Experiment 1 LDA Results

	Precision	Recall	F1-score
Non-seizure	0.86	0.96	0.91
seizure	0.74	0.41	0.53
Average	0.80	0.69	0.72

Precision refers to the proportion of correctly predicted positive observations among all positive observations, whereas recall/sensitivity refers to the fraction of correctly anticipated positive observations among all positive observations. The ratio between expected negative observations and actual negative observations is known as specificity, while the number of positive observations among the total observations is known as prevalence. The total number of observations properly predicted by the model is used to determine accuracy. The weighted average of the precision and recall scores is the F1 score.

3.3 Experiment 2

The model developed used LDA for feature extraction and a six-layer Artificial neural network for sorting out the dataset into seizure category and non-seizure category. Fig.9 present the Experiment 2 LDA Convergence Curve, while Fig.10 shows the Experiment 2 Model Accuracy.

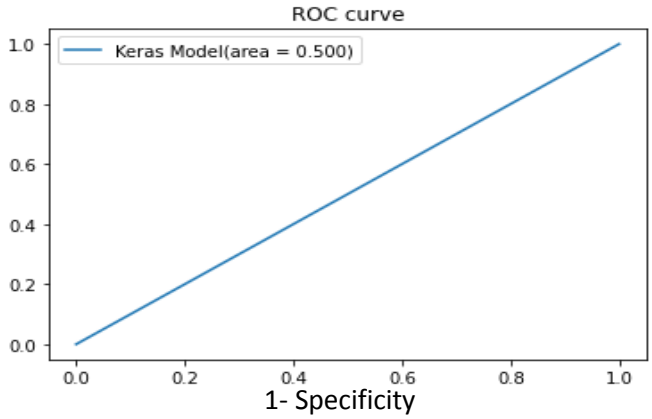

Fig.9 Experiment 2 LDA Convergence Curve

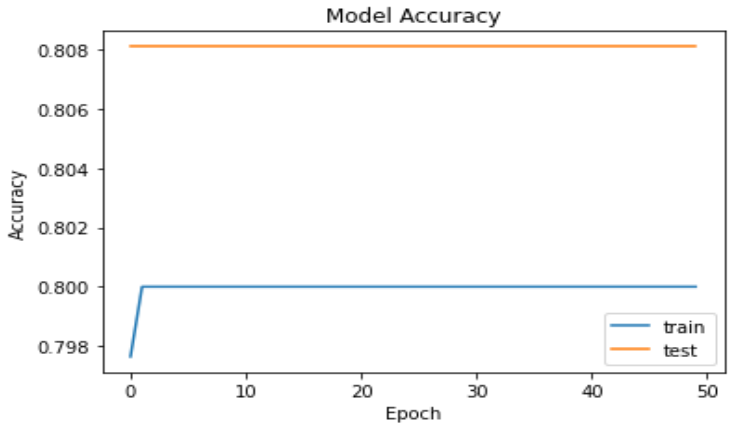

Fig.10 Experiment 2 Model Accuracy

Fig.9, shows the graph plots of the sensitivity against the value of 1- specificity or false positive rate of the ANN model. The ROC/ Convergence Curve has an AUC value of 0.5 with an accuracy of 80.3%. It can be observed that the true positive rate and also the false positive rate of the model increased at an average speed from 0 to 50 epochs. This suggests that the model is somewhat capable of capturing the key properties needed to accurately represent and differentiate seizure from non-seizure data. Fig.10 also, shows the accuracy curve against 50 iterations (epochs) of the model for both the training and test data. The training data has a fixed accuracy of 80% and the test data has an accuracy of 80.6%.

3.3.1 ANN Classification Results for Experiment 2

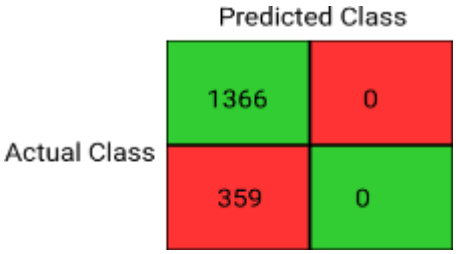

Fig.11 Experiment 2 LDA Confusion Matrix

As shown in Fig.11, the true positive value is zero, the true negative value is thirteen hundred sixty-six, the false positive value is zero, and the false-negative value is three hundred and fifty-nine. It can be observed that LDA correctly classifies 100 percent of non-seizure data while misclassifying 0 percent.

Again, it correctly classifies 100 percent of seizure data and misclassifies 0 percent of seizure data. The classification runtime ranged from 0.014s to 2s. This is to say that the system showed a fast classification time when the classification iterations were 50.

3.3.2 ANN Validation Results for Experiment 2

From the results as shown in Table-2, the overall model accuracy, sensitivity, precision, specificity and F1-score for experiment 2 is 80.3%, 50%, 40%, 100% and 46%, respectively.

Table-2 Summary of Experiment 2 LDA Results

	Precision	Recall	F1-score
Non-seizure	0.79	1.00	0.91
Seizure	0.00	0.00	0.00
Average	0.40	0.50	0.46

Precision refers to the proportion of correctly predicted positive observations among all positive observations, whereas recall/sensitivity refers to the fraction of correctly anticipated positive observations among all positive observations. The ratio between expected negative observations and actual negative observations is known as specificity, while the number of positive observations among the total observations is known as prevalence. The total number of observations properly predicted by the model is used to determine accuracy. The weighted average of the precision and recall scores is the F1 score.

3.4 Experiment 3

The model developed used LDA for feature extraction and a nine-layer Artificial neural network for separation into seizure and non-seizure categories.

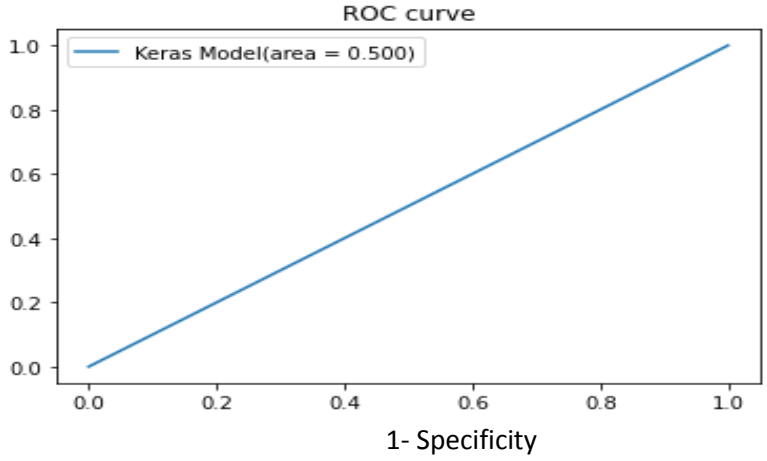

Fig.12 Experiment 3 LDA Convergence Curve

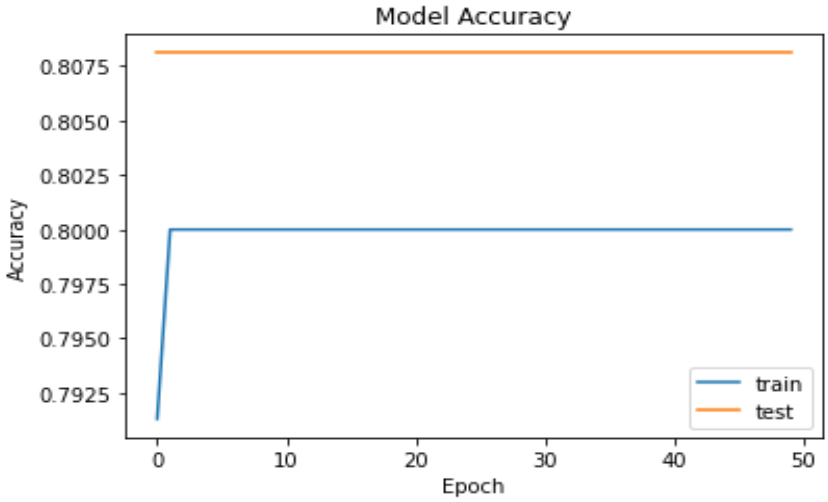

Fig.13 Experiment 3 Model Accuracy

Fig.12, shows a graph plot of the value of 1- specificity or false positive rate of the ANN model. The ROC/ Convergence Curve has an AUC value of 0.5 with an accuracy of 80.37%. It can be observed that the true positive rate and also the false positive rate of the model increased at an average speed from 0 to 50 epochs. This indicates that the model is somewhat capable of capturing the crucial properties needed to accurately represent and differentiate seizure and non-seizure data. Fig.13 also, shows the accuracy curve against 50 iterations (epochs) of the model for both the training and test data. The training data has a fixed accuracy of 80% and the test data has an accuracy of 80.75%. This is a result of the overfitting of the model.

3.4.1 ANN Classification Results for Experiment 3

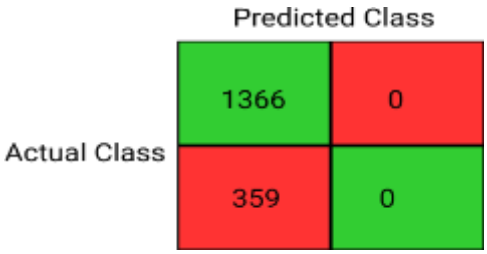

Fig.14 Experiment 3 LDA Confusion Matrix

As presented in Fig.14, the true positive value is zero, 1366 is the true negative value, 0 is the false positive value, and 359 is the false negative value. It can be observed that LDA correctly classifies 100 percent of non-seizure data while misclassifying 0 percent. Again, it correctly classifies 100 percent of seizure data and misclassifies 0 percent of seizure data. The classification runtime ranged from 0.025s to 2s. This is to say that the system showed a fast classification time when the classification iterations were 50.

3.4.2 ANN Validation Results for Experiment 3

From the results shows in Table-3, the overall model accuracy, sensitivity, precision, specificity and F1-score for experiment 1 is 80.37%, 50%, 40%, 100% and 44%, respectively

Table-3 Summary of Experiment 3 LDA Results

	Precision	Recall	F1-score
Non-seizure	0.79	1.00	0.88
Seizure	0.00	0.00	0.00
Average	0.40	0.50	0.44

Precision refers to the proportion of correctly predicted positive observations among all positive observations, whereas recall/sensitivity refers to the fraction of correctly anticipated positive observations among all positive observations. The ratio between expected negative observations and actual negative observations is known as specificity, while the number of positive observations among the total observations is known as prevalence. The total number of observations properly predicted by the model is used to determine accuracy. The weighted average of the precision and recall scores is the F1 score. The performed experiments show that the three-layer model in experiment 1 had the best performance. It had a general model accuracy, sensitivity, precision, specificity and an F1-score of 86%, 69%, 80%, 96.2% and 72%, respectively, with a classification time of 0.013s to 1s. The models in experiments 2 and 3 had similar results because the models began to overfit as more layers were added. The LDA-ANN model performs better when it has fewer layers.

3.5 Experiment 4

The model developed had a three-layer Artificial neural network for sorting out the dataset into seizure category and non-seizure category. Linear Discriminant Analysis was not used for the extraction of features in this experiment.

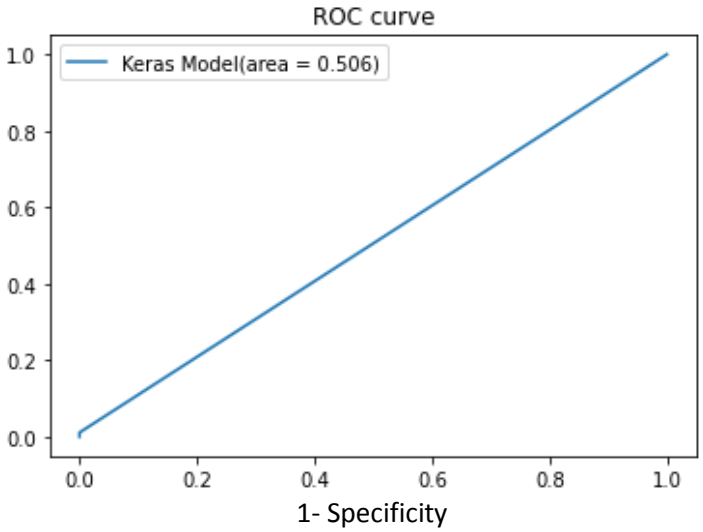

Fig.15 Experiment 4 ANN Convergence Curve

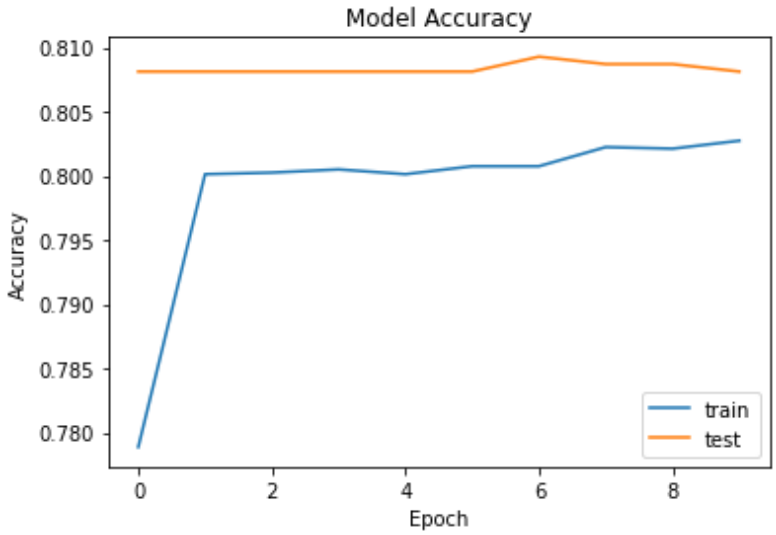

Fig.16 Experiment 4 Model Accuracy

Fig.15, shows a graph that plots the sensitivity against the value of 1- specificity or false positive rate of the ANN model. The ROC/ Convergence Curve has an AUC value of 0.51 with an accuracy of 80%. Fig.16 above shows the accuracy curve against 50 iterations (epochs) of the model for the training and test data. The accuracy of the training data increases from 0 to 80% for 0 to 1 epochs and has a general accuracy of 80%. The test data has a general accuracy of 81.0%.

3.5.1 ANN Classification Results for Experiment 4

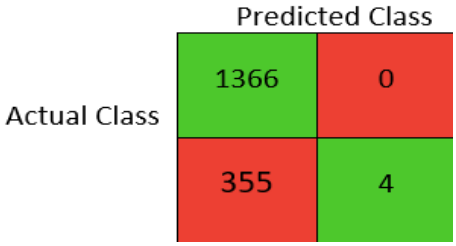

Fig.17 Experiment 4 ANN Confusion Matrix

As present in Fig.17, 4 is the true positive value, 1366 is the true negative value, 0 is the false positive value, and 355 is the false negative value. It can be observed that LDA correctly classifies 100 percent of non-seizure data while misclassifying 0 percent. Again, it correctly classifies 98.9% of seizure data and incorrectly classifies 1.1 percent of seizure data. The classification runtime ranged from 0.76s to 5s. This is to say that the system showed a slow classification time when the classification iterations were 50.

3.5.2 ANN Validation Results for Experiment 4

From Table-4, the overall model accuracy, sensitivity, precision, specificity and F1-score for experiment 4 is 81%, 50%, 90%, 100% and 46%, respectively.

Table-4 Summary of Experiment 4 Results

	Precision	Recall	F1-score
Non-seizure	0.79	1.00	0.89
Seizure	1.00	0.00	0.02
Average	0.90	0.50	0.46

Precision refers to the proportion of correctly predicted positive observations among all positive observations, whereas recall/sensitivity refers to the fraction of correctly anticipated positive observations among all positive observations. The ratio between expected negative observations and actual negative observations is known as specificity, while the number of positive observations among the total observations is known as prevalence. The total number of observations properly predicted by the model is used to determine accuracy. The weighted average of the precision and recall scores is the F1 score.

3.6 Comparison of Results

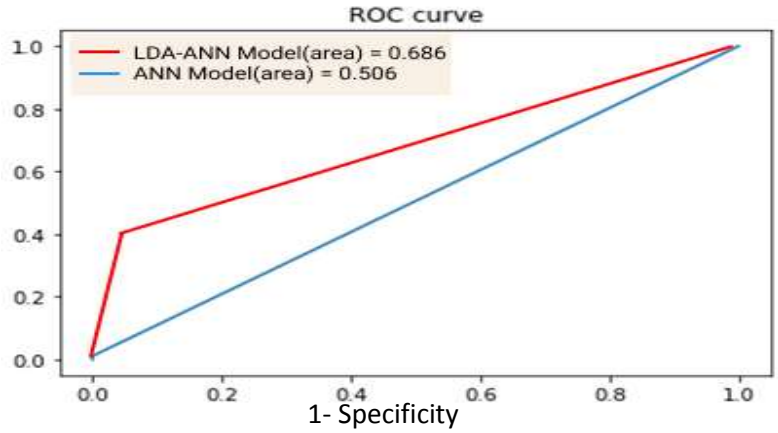

Fig.19 Convergence Curve of LDA-ANN and ANN

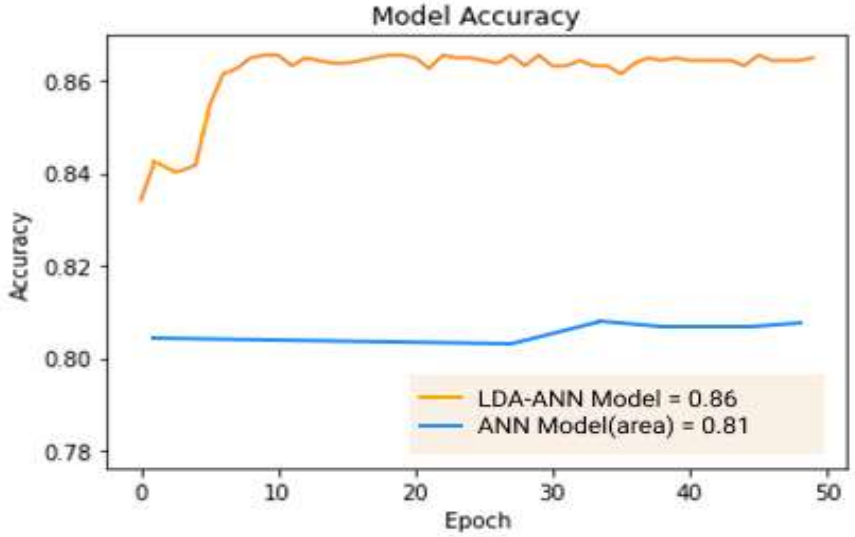

Fig.20 Model Accuracy of LDA-ANN and ANN

Fig.19 present the graph plots of the sensitivity against the value of 1- specificity or false positive rate of the LDA and ANN model. The ROC/ Convergence Curve for the LDA-ANN model has an

AUC value of 0.686 and a general accuracy of 86% while the ANN Model has an AUC value of 0.51 with a general accuracy of 80%.
Fig.20 present the accuracy curve against 50 iterations (epochs) of the model for both the LDA-ANN and ANN models. The general accuracy of the LDA-ANN model is 86% and the general accuracy of the ANN model is 81%. The LDA-ANN model has a 5% higher accuracy compared to the ANN model.

Table-5 Comparison of LDA-ANN and ANN models

Model	Precision	Recall	F1-score	Accuracy
LDA-ANN	0.80	0.69	0.72	0.86
ANN	0.90	0.50	0.46	0.81

From Table-5, the precision, recall, F1-score and general accuracy of the optimal LDA-ANN model is 80%, 69%, 72% and 86%, respectively. While, the precision, recall, F1-score and general accuracy of the ANN model is 90%, 50%, 46% and 81%, respectively.

RESEARCH CONTRIBUTION TO KNOWLEDGE

This study contributes to epilepsy classification by introducing smart and effective feature extraction with a low signal-to-noise ratio that extracts features using Linear Discriminant Analysis (LDA) and an Artificial Neural Network for classifying data into epileptic and non-epileptic seizure categories. Linear Discriminant Analysis is a technique for reducing dimensionality. It's utilized in Machine Learning and pattern categorization applications as a pre-processing phase. LDA is a supervised classification method used in the creation of competitive machine learning models. Before classification, linear discriminant analysis is performed to reduce the number of features to a more manageable quantity. Each of the additional dimensions is a template made up of a linear combination of pixel values. The benefit of LDA is that it takes information from both characteristics to construct a new axis, reducing variance and increasing class distance between the two variables.

CONCLUSION

Recurrent seizures are a symptom of epilepsy, which is a neurological disorder. These seizures can strike at any time and without warning. This influences the person's quality of life because they are unable to work or engage in routine activities without fear of having a seizure. Epilepsy can be diagnosed using an electroencephalogram (EEG). This study provides an intelligent LDA-ANN epilepsy prediction model that learns to identify data as "epileptic" or "non-epileptic" using a training dataset. After then, the model was used to forecast a new batch of data. The system aims to provide an accurate technique for classifying epilepsy seizures in patients, reduce human (physician) error, and shorten the time it takes to detect and predict seizures in patients. Because this technology may be deployed to a real-time system and made available to the public, this research also tackles the issue of neurophysiologists being scarce in poor countries. This method was put into place and demonstrated with the help of LDA and ANN dependencies and Anaconda software. The LDA-ANN system was evaluated using the accuracy, sensitivity, specificity, precision and F1-score. The developed system has an overall accuracy, sensitivity, precision, specificity and F1-score of 86%, 69%, 80%, 96.2% and 72%, respectively, with a response time of 0.013s to 1s.

RECOMMENDATION

For an intelligent and real-time prediction of epileptic seizures, portable real-time implementation of the system is recommended. This would forecast seizures in real-time, allowing the patient to prepare for them. This can aid in seizure prevention and effective seizure treatment. This would considerably improve the patient's quality of life while also reducing the number of accidents caused by seizures.

CONFLICT OF INTEREST

This research was conducted by us, and there are no conflicts of interest involved.

ACKNOWLEDGEMENT

This work was supported in part by FUTMINNA under NEEDs Assessment Grant 2017.

REFERENCES

Acharya, U. R., Hagiwara, Y., & Adeli, H. (2018). Automated seizure prediction. *In Epilepsy and Behavior* (Vol. 88, pp. 251–261). Academic Press Inc.

Amin, S. U., Alsulaiman, M., & Muhammad, G. (2019). Applying Deep Learning for Epilepsy Seizure Detection and Brain Mapping Visualization. 15(1), pp. 1–17

Daoud, H., & Bayoumi, M. (2019). Efficient Epileptic Seizure Prediction based on Deep Learning. *IEEE Transactions on Biomedical Circuits and Systems*, pp.(c), 1.

Falco-Walter, J. J., Scheffer, I. E., & Fisher, R. S. (2018). The new definition and classification of seizures and epilepsy. *Epilepsy Research*, 139(July 2017), 73–79.

Kiral-Kornek, I., Roy, S., Nurse, E., Mashford, B., Karoly, P., Carroll, T., Payne, D., Saha, S., Baldassano, S., O'Brien, T., Grayden, D., Cook, M., Freestone, D., & Harrer, S. (2018). Epileptic Seizure Prediction Using Big Data and Deep Learning: Toward a Mobile System. *EBioMedicine*, 27, pp. 103–111.

Lasefr, Z., Shiva, S., Ayyalasomayajula, V. N. R., & Elleithy, K. (2017). Epilepsy Seizure Detection Using EEG signals, pp. 162–167.

Liu, J., & Woodson, B. (2019). Deep Learning Classification for Epilepsy Detection Using a Single Channel Electroencephalography (EEG).

Maria, H., Heller, S., Watter, M., Blum, M., Manzouri, F., Matthias, D., Schulze-bonhage, A., Woias, P., & Boedecker, J. (2018). Early Seizure Detection with an Energy-Efficient Convolutional Neural Network on an Implantable Microcontroller.

Murugavel, S. R. A. S. M. (2018). Epileptic seizure detection using fuzzy - rules - based sub-band-specific features and layered multi-class SVM. *Pattern Analysis and Applications*, 0123456789, pp. 15–18.

Sharmila, A., & Mahalakshmi, P. (2017). Wavelet-based feature extraction for the classification of the epileptic seizure EEG signal. *Journal of Medical Engineering & Technology*, pp.1–11.

Shanir, P. P. M., Khan, K. A., Khan, Y. U., Farooq, O., & Adeli, H. (2017). Automatic Seizure Detection Based on Morphological Features Using One- Dimensional Local Binary Pattern on Long-Term EEG.

Varnado, S., & Price, D. (2020). Basics of modern epilepsy classification and terminology. Current Problems in Pediatric and Adolescent Health Care, 50(11), 100891.

World Health Organization. (2019). WHO | Epilepsy: a public health imperative. In Who. https://www.who.int/mental_health/neurology/epilepsy/report_2019/en/