ARTICLE

Analysis of EEG signals for the detection of epileptic seizures using feature extraction

Ravi Kumar Kandagatla*[1], V. Jayachandra Naidu[2], PS. Sreenivas Reddy[3], Veera Kavya Pandilla[1], Marriwada Joshitha[1], Chanamala Rakesh[1]

[1]Department of Electronics and Communication Engineering, Lakireddy Bali Reddy College of Engineering (Autonomous), Mylavaram-521230, Andhra Pradesh, India
[2]Department of Electronics and Communication Engineering, Sri Venkateswara College of Engineering & Technology (Autonomous), Chittoor, India
[3]Department of Electronics and Communication Engineering, Nalla Narasimha Reddy Education Society's group of Institutions, Telangana, India

ABSTRACT

The electroencephalogram, which tracks electrical signals in the central nervous system, has been extensively used to diagnose epilepsy, which represents a particular sort of brain abnormality. However, developing seizure classification techniques with significantly better precision and reduced complexity remains challenging. The Epileptic Seizure Recognition dataset, which is publicly accessible in the Kaagle and in the machine learning repository, was used to identify seizures. To identify the seizure, we compared six classification methods to determine which one had the highest success rate. The dataset is subsequently divided, trained, and tested in order to categorize it further using six machine learning algorithms: Stochastic Gradient Descent, Logistic Regression, Naïve Bayes, K-Nearest Neighbors Algorithm, Extra Tree Classifier and Decision Tree. When contrasted with alternative techniques, Extra Trees Classifier possesses the highest accuracy results. The algorithm attained a 96 percent success rate.

NOMENCLATURE

EEG-Electroencephalogram
MRI - Magnetic Resonance Imaging
PET- Positron Emission Tomography
DWT- Discrete Wavelet Transform
TFA-Time Frequency Analysis
STFT -Short-Time Fourier Transform
WT -Wavelet Transform
PSD- Power spectral density
FFT-Fast Fourier Transform
KNN- K-Nearest Neighbors
SGD- Stochastic Gradient Descent
NB- Naïve Bayes

*Corresponding author
Email Address: 2k6ravi@gmail.com
Date received: May 6, 2023
Date revised: December 30, 2023
Date accepted: December 31, 2023
DOI: https://doi.org/10.54645/202417SupCQP-72

KEYWORDS

Epileptic Seizure, Feature Extraction, Machine Learning, Classification, Electrical Signals, Stochastic Gradient Descent, K-Nearest Neighbors Algorithm, Naive Bayes, Decision Tree, Logistic Regression, Extra Tree Classifier, Accuracy.

DT- Decision Tree
ETC- Extra Trees Classifier
TP -True Positive
TN- True Negative
FP- False Positive
FN- False Negative
ROC-Receiver Operating Characteristics
AUC- Area Under the Curve

INTRODUCTION

The field of neuroscience focuses on the nervous system study which includes the brain, peripheral nerves and spinal cord. Neuroscience is a multidisciplinary field that combines psychology, biology and other disciplines to understand the function of nervous system, how it relates to behavior, cognition, and emotion. One area of interest within neuroscience is the study of neurological disorders, (Quanhong Wang et al. 2023) which can affect the nervous system's function and lead to various symptoms and disabilities. Different ranges of factors are responsible for disorders which include genetics, environmental factors, infections, or injuries (Burgess PW 2016). Some common neurological disorders are Parkinson's disease, Alzheimer's disease, epilepsy and multiple sclerosis. Epilepsy is a neurological disorder which affects several thousands of people worldwide, characterized by recurrent seizures that can be disabling and potentially life-threatening (Bajpai R et al. 2022). Seizures can manifest in various forms, making diagnosis and treatment a challenging task for medical professionals (Anca-Mihaela Vasilica et al.2023). Various tools are developed for proper diagnosis and treating of neurological disorders. This happens due to technological advancement in neuroscience field (Kumar TS et al. 2015). For instance, the imaging technique like magnetic resonance imaging (MRI) is helpful for understanding the working of brain and positron emission tomography (PET), which helps in visualizing the structure of brain and its function (Loukas Ilias et al. 2023). Moreover, contemporary medical practices leverage artificial intelligence and machine learning algorithms to enhance diagnostic accuracy. These techniques (Yang Li et al. 2016) utilize the training and testing data for efficient feature extraction. Medical data related to electroencephalography (EEG) recordings, accelerometers, and video cameras, are processed to identify patterns related to seizure condition and non-seizure condition of patients (Usman SM et al. 2020). Epilepsy is a serious and vastly spreading neurological disorder that is affecting around 65 million people worldwide, and impacts about one in every 26 people at some point in their lives. Various types of seizures, which are characterized by symptoms like jerk in movements, loss of consciousness that may be temporary or permanent and confusion behavior are seen in people. There are types of seizures which may not be visible, but patients may display symptoms such as staring without any thought for a short period. Most of the seizures can occur unpredictably, and hence patients may suffer with injuries due to sudden falls, tongue biting, or loss of bladder or bowel control. Thus, detecting seizures is crucial for individuals under medical surveillance who are at risk of seizures (Li Y et al. 2012). This work aims to develop a classification system to predict whether a person is having a seizure or not having seizure using binary

classification. The classification system will use machine learning techniques to analyze data related to the patient's physiological and behavioral characteristics during the seizure. The data will be collected from various sources, such as EEG recordings, accelerometer sensors, and video recordings. Different patterns in the data are analyzed using machine learning algorithms by providing training. Training helps to identify patterns in the data which distinguishes between seizure and non-seizure states (Olmi B et al. 2021). The proposed classification system will have significant clinical implications, as it can aid medical professionals in the early detection and treatment of seizures. For instance, doctors can use the system to alert caregivers or family members when a seizure is about to occur, allowing for prompt intervention. The system can also help medical professionals track the patient's seizure frequency and response to treatment, providing valuable information for personalized care. Overall, the study of neuroscience and neurological disorders has significant implications for understanding how the nervous system functions and how it relates to behavior, cognition, and emotion. Furthermore, advances in research in neuroscience field and technological advancements leads to new diagnosis techniques and new tools for treating and diagnosis of neurological disorders, improving patient outcomes and quality of life (Kai Fu et al. 2014). In this work, we aim to develop a classification system to predict whether a person is experiencing epilepsy behavior / seizure or not using machine learning algorithms. We will collect data from various sources, such as EEG recordings, accelerometers, and video cameras, to train and test the classification system's accuracy and reliability (Ghosh Dastidar S et al. 2007). The proposed system will have significant clinical implications, as it can aid medical professionals in the early detection.

METHODOLOGY

Dataset Description

The UCI machine learning repository or Kaagle provides access to a dataset that contains 4097 electroencephalograms (EEG) readings per patient, collected over a 23.5 second period, with a total of 500 patients. In order to transform the data into a suitable format for analysis, the 4097 readings were divided equally into 23 chunks per patient, resulting in 23 rows of data per patient. Each row of data represents one second of EEG readings and contains 178 columns of data. These columns represent the individual readings taken during that second. Overall, the dataset comprises eleven thousand five hundred rows of data and hundred and eighty columns, with the first column containing the patient ID and the last column indicating the patient's status, i.e., whether the person is experiencing a seizure or not (Shu Lih Oh et al. 2023). This dataset is particularly useful for developing and testing algorithms to detect seizure activity from EEG signals. By breaking the data down into one-second chunks, researchers can analyze the EEG signals in a more granular manner, potentially enabling more accurate detection of seizure activity. Additionally, the large number of patients and readings provides a robust dataset (Liu et al. 2016) that can be used to validate the effectiveness of the developed algorithms (Shu Lih Oh et al. 2023). The dataset consists of several columns containing important information for seizure detection. One of these columns is a hashed patient ID that ensures the privacy of the patient's personal information. Another crucial column contains 178 electroencephalogram (EEG) readings, which represent one second of EEG data for the patient. The final column contains an output variable, "y" which indicates the patient status during that particular monitoring second. If the patient is experiencing a seizure, "y" is assigned the value of 1, while all other numbers represent different statuses that are not relevant to seizure detection. To effectively detect seizures, the "y" output variable is converted into two possible outputs as binary (two) variable, which turns this detection problem into a classification task. In this way, the dataset can be used to train different models using machine learning algorithms to

accurately distinguish between seizure suffering patient and non- seizure suffering patient based on the EEG readings. By leveraging the information contained in the EEG data, machine learning models are trained and learned for features extraction and pattern identification that are indicative of seizure activity, thereby helping to diagnose and treat epilepsy more effectively.

Implementation Model

Implementation and procedure for the seizure detection is given in (Figure 1 and Figure 2)

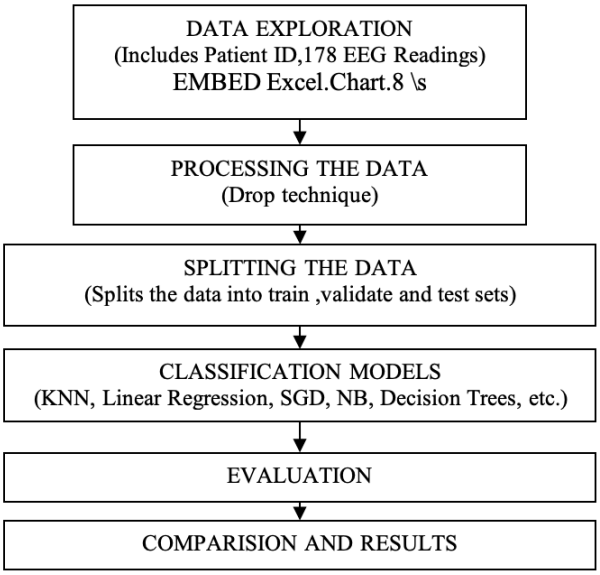

Figure 1: Steps in proposed Method

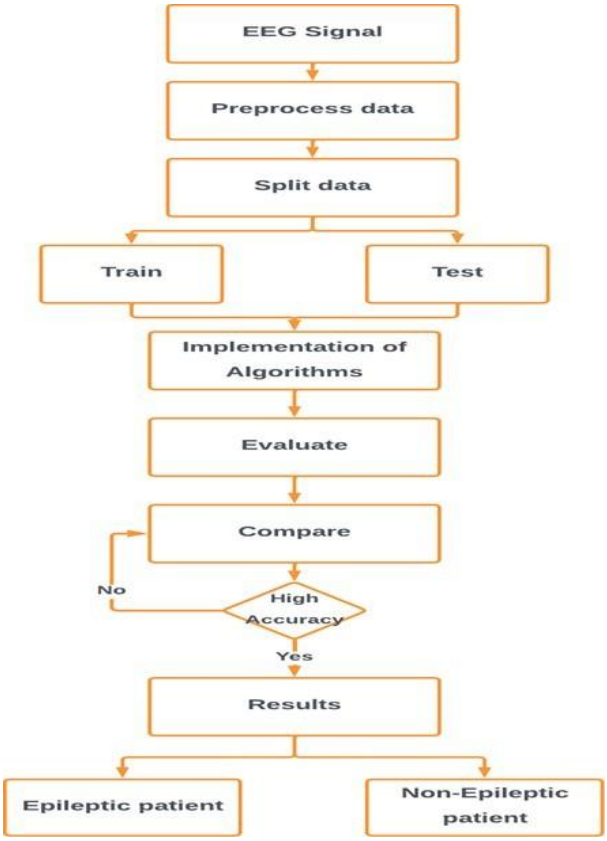

Figure 2: Implementation block diagram

Processing The Data

The data is preprocessed using a technique drop to remove the unwanted columns in the dataset (Figure 1). It creates a new column called "OUTPUT LABEL" which is a binary variable based on patient suffering from seizure or not (y == 1). The next line of code converts the "OUTPUT LABEL" column to integer values. The third line of code removes the "y" column from the dataframe using the pop() method, and the last line drops the first column of the data using the drop() function. Here the parameter is set to 1 which indicates that it is a column that needs to be removed (Figure 2).

Splitting the dataset

Firstly the data gets checked to find any duplicates present in the columns in order to split them into train, validate and test sets. The process performs two checks to ensure the integrity of the dataset. The first check ensures that there are no duplicate columns in the DataFrame by creating a list of column names and checking for duplicates. If there are duplicates, the assert statement will throw an error message indicating that there are du- plicate columns found in the DataFrame. The second check verifies that the total number of columns in the DataFrame is one more than the length of the cols input list. This is because the dataset should have one additional column for the target variable that was added during preprocessing. If there is a mismatch in dimensions between cols input and the DataFrame, the assert statement will throw an error message indicating a mismatch in dimensions of df data or cols input. This check is important as it ensures that the dataset is in the correct format for further analysis and modelling. Step1: To ensure that there is no order or pattern in the samples of the dataset, the first step is to shuffle the dataset (Figure 1). Training set size generally depends on the dataset and the number of samples present in that dataset, with the training set ranging from 50 percent to 90 percent of the original dataset. In general, the more the number of samples, the higher the percentage of the dataset that can be assigned to the training set. The size of the validation and testing sets are considered as same size. Step2: The dataset will be split into three sets – training set, validation set, and testing set - using a 70 training/15 training/15 training split ratio. The first step is to separate the validation and testing sets from the training set to ensure that all three sets have similar distributions. Step3: The subsequent step is to balance our dataset to prevent the creation of a model that wrongly categorizes samples as belonging to the majority class. This indicates the patients not experiencing a seizure.

Classification Models

In this approach, the default hyper-parameters for all the models will be used, and the model with the highest validation score will be selected. This approach is called the baseline approach, where we use the default hyper-parameters to create all the models. Once all the models are created, the one with the highest score is selected for validation, and its performance is compared against other models. The advantage of using this approach is that it is a quick way to compare multiple models and select the best one without spending too much time tuning each model's hyper-parameters. However, this approach may not always lead to the best model, as some models may require specific hyper-parameters to achieve their best performance. Therefore, after selecting the best model, it is essential to fine-tune its hyper-parameters to optimize its performance further.

KNN

The K-Nearest Neighbors (KNN) model is a basic binary classification learning model in machine learning. This model assigns classes to input samples based on the number of nearest samples that are most like it. For instance, if we set k as 3 and all three of the nearest samples are from the positive class, then the input sample will also be classified as a positive class.

However, if only two out of three nearest samples belong to the positive class, then the input sample will have a probability of 66. We are training and evaluating a KNN (K-Nearest Neighbors) classification model for seizure prediction on both the training and validation sets. The model is first fitted using the data allocate for training and then used to make predictions on the training set as well as validation sets. We then evaluate the performance of the model on both sets using several evaluation metrics such as AUC (Area Under the Curve), accuracy, recall, precision, specificity, and prevalence. These metrics give us an idea of how well the model is performing in terms of correctly identifying both positive and negative cases (Mingkan Shen et al. 2023). The performance of the model on the training set and validation set is printed separately, along with the values of each evaluation metric. The goal is to select the model with the best performance on the validation set to use for making predictions on new, unseen data.

Naive Bayes

The naive Bayes classifier is uses Bayes theorem from probability concept to predict the probability of a sample belonging to a certain class. This model works by assuming that all features are independent of each other, meaning the probability of seeing a set of features is simply the product of their individual probabilities. However, this assumption is often unrealistic since features may have some correlation or interaction with each other. As a result, the model may not perform well in certain situations. The Naive Bayes classifier using Gaussian distribution uses the Bayes theorem to classify the sample based on the probability of having the features given the two classes positive and negative. This technique assumes that all available features are independent and mutually exclusive to each other. However, this is not true in practice. We fit the model on the training set and used it to predict on the training and validation sets. We calculated the AUC, accuracy, recall, precision, and specificity for both sets. These metrics allow us to evaluate how well the model is performing in different aspects. Metrics are useful false positives and for correctly identifying the seizure positives. The evaluation using naïve bayes shows the performance of the model on the two sets, which allows us to determine the overfitting nature of model to the training data or if it can generalize well to new data.

Decision Tree

A decision tree model determines the class of a sample by asking a series of questions that divide the data of same class into different sub-regions. The model continues to ask these questions until all the samples are sorted into pure categories or until it meets a certain criterion. However, decision trees are not highly accurate and typically only perform marginally better than randomly guessing. Additionally, these models are prone to overfitting the training data. A decision tree model determines the class of a sample by asking a series of questions that divide the data into same class sub-regions. The model continues to ask these questions until all the samples are sorted into pure categories or until it meets a certain criterion. However, decision trees are not highly accurate and typically only perform marginally better than randomly guessing. Additionally, these models are prone to overfitting the training data. The decision tree classifier is trained on the training data and the predict_proba ( ) method is used to obtain probability estimates for the positive class. These probability estimates are used to compute several evaluation metrics such as AUC, accuracy, precision, recall and specificity on both the training and validation data.

Logistic Regression

Logistic regression is a basic logistic model that is an extension of regular Logistic models. Unlike Logistic models, it predicts binary outcomes, i.e., whether something is true or false. Logistic regression models estimate a Logistic decision boundary for both classes, which is then transformed into a probability using a sigmoid function. This model is effective when there is a clear separation between the negative class and positive class in the data. It is essential to scale all the features and ensure that the dependent variable is binary for this model to work effectively. The model analyses and predicts the condition for true or false. It does this by fitting a straight line that separates the positive and negative class in the data. The line is passed through a function called the sigmoid function to convert the Logistic decision boundary into probabilities of the sample belonging to the positive class. Logistic regression performs well when there is a clear separation between the classes in the data. This model scales the available features and that the dependent variable has only two possible outcomes. We fit the model to the training data and use it to predict probabilities of the samples in the validation set belonging to the positive class. We then use these probabilities to evaluate the model performance on the training set as well as validation set.

SGD

Gradient descent is a popular algorithm used to minimize the loss function in various models such as Logistic regression and clustering models. Its aim is to optimize the Logistic function in these models. Stochastic gradient descent is a variation of gradient descent that allows mini-batch learning. Instead of using the whole dataset to take a step, SGD uses multiple samples to do so. This method is particularly useful when there is duplicate or redundancy in the data, which is often seen in clustering. So, the SGD classifier works faster than to available models. The SGD Classifier is an algorithm used to minimize many loss functions in various models, such as Logistic and logistic regression. It is like logistic regression in that it optimizes a Logistic function. However, the difference is that SGD classifier uses minimum sized data for learning, where the model uses a subset of the data to proceed further instead of the entire dataset. This is more useful in cases where there is redundancy in the data, which is often observed in clustering. Because of its ability to use mini-batch learning, SGD Classifier is much faster than traditional logistic regression.

Random Forest

Random forest was developed as a solution to the over-fitting issues that decision trees tend to encounter. Random forest comprises multiple decision trees that work together as a model. Each tree is constructed using features selected from sub-set selected randomly and a bootstrapped dataset. This process reduces the correlation between the trees, thus decreasing the likelihood of overfitting. Additionally, we can evaluate the performance of this classifier using the “out-of-bag” data that wasn’t used in any of the trees. Compared to decision trees, random forest has lower variance and is often preferred since it provides better generalization.

Extra Trees Classifier

The ExtraTrees Classifier is a variant of the popular Random Forest model. However, there are some differences between them. For example, when deciding which variable to use to split a node, the ExtraTrees Classifier randomly selects samples from the give training data set than using bootstrapped samples like the Random Forest. Additionally, instead of specifying node splits, the ExtraTrees Classifier chooses them randomly. These differences make the Extra Trees Classifier less likely to overfit, and often results in a simple model than

the Random Forest. We are using the Extra Trees Classifier to classify our data. The Extra Trees Classifier is an average model that combines various decision trees to reduce overfitting. We first define our Extra Trees Classifier with certain parameters like, number of decision trees in the forest and the criteria for splitting nodes. Then we fit the model on our training data. We then use the trained model to predict probabilities of the positive class for both the training and validation sets (Li Y et al. 2012). These predicted probabilities are used for calculation of various performance measures like accuracy, recall, and precision. Finally, we print out the results of these performance metrics for both the training and validation sets. This allows us to evaluate how well our Extra Trees Classifier is performing and make any necessary adjustments to the model or hyper- parameters.

This work focused on developing a machine learning model that could classify EEG readings to determine whether a patient was having a seizure or not. Seizures are a medical emergency (Boashash B et al. 2012) , and their prompt recognition and treatment are crucial to prevent long-term health consequences or even death. Therefore, having an accurate and reliable method for detecting seizures is of utmost importance. The K-Nearest Neighbors (KNN) model has achieved an AUC of 0.992 and an accuracy of 0.626 on the training set, which suggests that the model has learned to distinguish between the two classes reasonably well. However, the recall score is low, indicating that the model may have difficulty correctly identifying positive cases (i.e., it has a high false negative rate). It is observed that, it provides high precision score which indicates that the model is likely to produce few false positives. When evaluated on the validation set, the KNN model achieves an AUC of 0.966 and an accuracy of 0.852, suggesting that it generalizes reasonably well to new data. However, the recall score is still relatively low, which indicates that the model may struggle to identify all positive cases in new data. It is worth noting that the prevalence of positive cases is relatively low in the training (0.5) and validation (0.202) sets (0.5 and 0.202, respectively), which may have affected the performance of the KNN model. The Naive Bayes model has performed quite well on the training as well as validation sets, with AUC scores of 0.983 and 0.986, respectively. The accuracy score is also high, indicating that the model is predicting correctly for most of the data points. The recall score is also high, indicating that the model is able to identify a high proportion of the positive cases. The precision score is also high, indicating that the model is correctly identifying a high proportion of the predicted positive cases. The high specificity score is indicating that the model can identify negative cases.

The Decision Tree model is providing high accuracy on the training set (98.2 percent) but performed worse on the validation set (89.7 percent), indicating some degree of overfitting. The AUC on the training set was 0.985 and on the validation set was 0.843. The recall on the training set was high (96.4 percent) but the precision was perfect (99.9 percent), indicating potential overfitting. The Logistic Regression model had a low AUC score on the training set (0.623) and on the validation set (0.512), indicating poor discrimination between classes. The accuracy was moderate on the training set (64.6 percent) and validation set (69.3 percent), and the recall and precision were also low. The SGDC model performed similarly to the Logistic Regression model, with a low AUC score on both the training sets (0.569) and validation set (0.486), and moderate accuracy on both sets (58.9 percent on training and 61.1 percent on validation). The Random Forest model performed very well on the training set, with high accuracy (96.9 percent), recall (94.5 percent), and precision (99.2 percent). However, there was some degree of overfitting, with a drop in performance on the validation set (AUC of 0.992 and accuracy of 96.0 percent). The Gradient Boosting Classifier achieved perfect performance on the training set, with an AUC of 1.0, accuracy of 100 percent, recall of 100 percent, and precision of 100 percent. However, there was some degree of overfitting, as the performance dropped on the validation set (AUC of 0.99 and accuracy of 95.5 percent). The Extra Trees Classifier is a type of ensemble model that works by various ensemble of decision trees and combining their results to make predictions. It seems to perform very well on the training and validation sets, with high AUC values and accuracy scores above 0.96. On the training set, it achieves a perfect AUC of 1.0, indicating that it can separate the positive and negative cases perfectly. On the validation set, the model's AUC score is still very high at 0.995, indicating good performance on unseen data. Recall metric measures positive cases. The ETC model has high scores for both the training and validation sets. The precision scores, which measure the proportion of true positive cases out of all predicted positive cases, are also relatively high for both sets. The ETC model seems to be performing very well on this dataset, suggesting that it may be a good choice for further analysis. For predictions with the highest accuracy, our algorithm opted for ETC as the best classifier.

RESULTS AND DISCUSSION

To evaluate the performance of our classification models, we decided to use the AUC curve metric, which is commonly used as it is not affected by the threshold chosen. This allows for a fair comparison between models (Figure 3). After plotting the AUC curves of all six models, we noticed that four of them performed exceptionally well, which can be attributed to the stark differences in EEG readings of seizure and non-seizure patients. However, the decision tree model showed signs of overfitting, as there was a noticeable gap between the training and validation AUC curves. Plotting the AUC values for a specific machine learning model over a range of different values is used for the maximum number of features in the model (Figure 4). The graph has two lines, one for the AUC values on the training data and another for the AUC values on the validation data. The x-axis of the graph represents the maximum number of features used in the model. The y-axis represents the AUC value, which is a measure of the model's performance. A higher AUC value shows that the model is providing better performance in distinguishing between the positive class and negative class. The graph is titled "Effect of Max Features on AUC," indicating that it shows how the performance of the model changes as the maximum number of features is varied. The labels on the x and y axes clarify the information shown in the graph. Based on the graph, it appears that the model's performance on both the training and validation data improves as the maximum number of features increases, up to a point where it plateaus. This suggests that increasing the number of features beyond this point may not result in significant improvements in the model's performance.

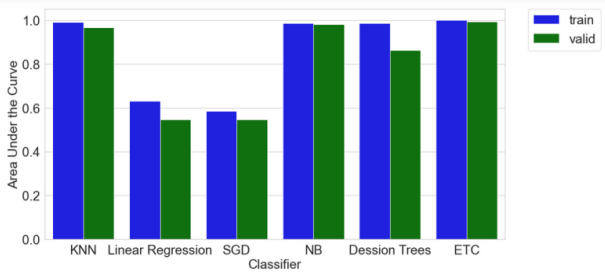

Figure 3: Comparison of ML Models

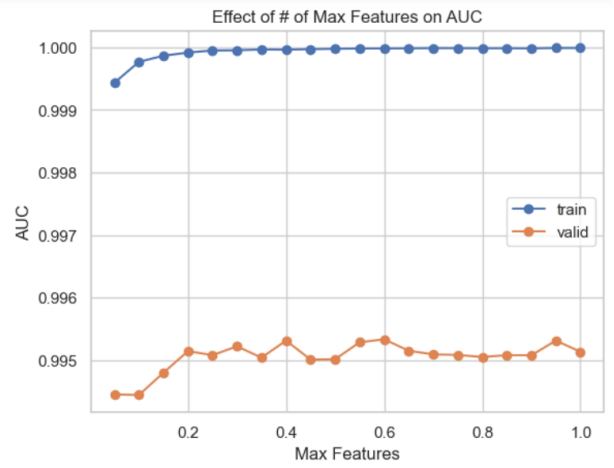

Figure 4: Learning Curve

Confusion Matrix :

There are two confusion matrices for the training and validation datasets. Confusion matrix is a matrix representation as table that is used for evaluating the performance of a various classification models. It summarizes the actual and predicted classes, highlighting how many observations were classified correctly and incorrectly by the model (Figure 5). The training confusion matrix indicates the number of true negative (TN), false negative (FN), true positive (TP) and false positive (FP), and predictions made by the model on the training data (Figure 5). In this case, the confusion matrix has two rows and two columns, corresponding to the two possible classes (seizure and non-seizure).The 0th row of the confusion matrix represents the true negative (TN) predictions made by the model, which means the model correctly predicted that the EEG readings did not indicate a seizure. The 0th column of the 0th row corresponds to the TN predictions when the actual class is negative (non-seizure), while the 1st column of the 0th row corresponds to the TN predictions when the actual class is positive (seizure).In the given confusion matrix, the value of 1.6e+03 in the 0th column of the 0th row indicates that the model made 1,600 TN predictions when the actual class was negative (non-seizure), while the value of 0 in the 1st column of the 0th row indicates that the model did not make any TN predictions when the actual class was positive (seizure). This means that the model did not predict any false negatives (FN) in the training data, indicating a high sensitivity of the model to detect seizures. In the validation confusion matrix, the 0th row represents the true negative (TN) and true positive (TP) classifications, while the 1st row represents the false negative (FN) and false positive (FP) classifications. The value of 1.3e+03 in the 0th column of the 0th row indicates that there were 1.3e+03 instances where the model correctly classified a seizure-free patient as not having a seizure (TN). The value of 50 in the 1st column of the 0th row indicates that there were 50 instances where the model incorrectly classified a seizure- free patient as having a seizure (FP). The occurrences in row and column 0 represent the number of instances where the model predicted a seizure-free patient, and the true label was also seizure-free. In this case, the model correctly predicted the label, so the value is 5. Similarly, the value of 3.4e+02 in row 1 and column 1 indicates that the model correctly classified 340 patients who were having a seizure (TP.) In summary, the confusion matrix provides information on how well the model is performing in terms of different types of classifications, and the values in each cell provide insight into the number of instances where the model correctly or incorrectly classified patients with or without seizures. The y-axis represents the true positive rate, or the proportion of actual seizure events that are correctly classified as seizures. Ideally, the ROC curve should be positioned as close as possible to the upper left corner of the plot, which represents a classifier with perfect discrimination between seizure and non-seizure events. The area under the curve (AUC) is a measure of the overall performance of the classifier, with an AUC of 1 indicating perfect discrimination and an AUC of 0.5 indicating a classifier that performs no better than random guessing. In the case of seizure detection, a high true positive rate is crucial to accurately detect seizure events, while the false positive rate should be minimized to avoid misclassifying non-seizure events as seizures.

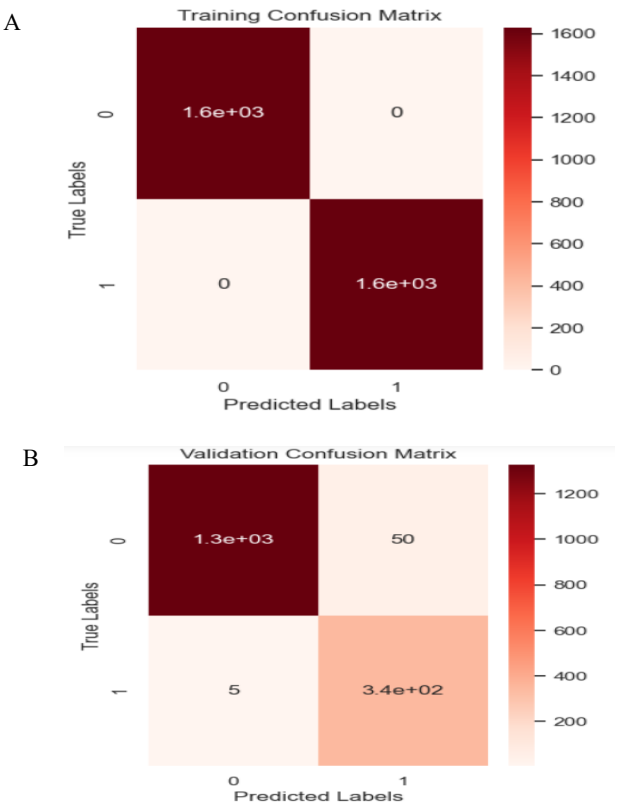

Figure 5: Confusion Matrix (A) Training (B) Validation

ROC Curve

Therefore, the ROC curve can help to determine the optimal threshold for the binary classifier that balances these two objectives.

A ROC (Receiver Operating Characteristic) curve is a graphical representation of the performance of a binary classifier system as its discrimination threshold is varied. It shows the trade-off between the true positive rate (sensitivity) and the false positive rate (1-specificity) as the threshold is changed. In the context of seizure detection, the ROC curve can be used to evaluate the performance of a binary classifier that attempts to detect seizures based on EEG (Electroencephalogram) readings. The x-axis of the ROC curve represents the false positive rate, or the proportion of non-seizure events that are incorrectly classified as seizures (Figure 6).

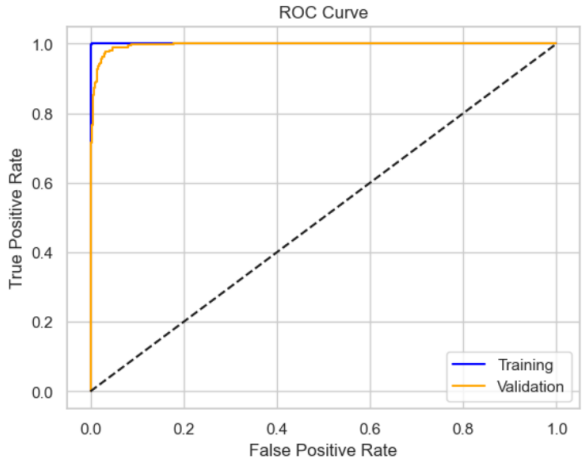

Figure 6: ROC curve

CONCLUSION

To achieve the goal, we employed several machine learning models, including Decision Tree, Logistic Regression, Stochastic Gradient Descent Classifier, Random Forest, Gradient Boosting, and Extra Trees Classifier. We trained these models on a dataset consisting of electroencephalogram readings from patients who had previously experienced seizures and those who had not. The dataset had 5,000 instances, with 2,523 belonging to the positive class (i.e., seizure occurrences) and 2,477 to the negative class (i.e., no seizure occurrence). After training the models, we evaluated their performance using several metrics, including lift, accuracy, recall, precision, specificity, and prevalence. Lift measures the effectiveness of the model in comparison to random guessing, while accuracy indicates how often the model correctly predicts the outcome. Recall measures the proportion of actual positive instances that the model identifies correctly, while precision measures the proportion of instances that the model identifies as positive that are indeed positive. Specificity indicates the measure of negative proportion instances that the model identifies correctly, while prevalence indicates the part of the positive class. In particular, the Extra Tree Classifier model achieved an Area Under the Curve of 0.995, indicating that it has excellent discrimination power between the positive and negative classes. Moreover, the model had a recall of 0.966, indicating that it correctly identified 96.6 percent of patients experiencing seizures. The precision was also high, with the model identifying 87.5 percent of the instances identified as positive being positive. This indicates that the model has a less false positive rate and can accurately identify patients experiencing seizures. The successful performance of the Extra Tree Classifier model can be attributed to its ability to generate many random decision trees and aggregate their outputs to make predictions. This approach results in reduced variance and overfitting, leading to improved performance on test datasets. The Extra Tree Classifier model's high accuracy, recall, and precision indicate that it is a suitable choice for real-world applications of seizure detection. In conclusion, using electroencephalogram measurements and machine learning, we have effectively created a classification model that can predict the likelihood that a patient will experience seizures. With a lift metric of 4.3 and an accuracy of 97.4 percent in predicting positive classes in the test dataset, the ExtraTree Classifier model outperformed the other evaluated models as shown in Table 1. Due to its great precision and recall, it can accurately identify patients experiencing seizures while maintaining a low rate of false positives. The effectiveness of this model offers a potential basis for creating trustworthy seizure detection technologies that can enhance patient outcomes and save lives.

Table 1: Training and Validation Results

Classifier	Performance Measure	Training	Validation
KNN	AUC	0.992	0.966
	Accuracy	0.626	0.855
	Recall	0.253	0.257
	Precision	0.998	0.989
	Specificity	0.999	0.999
	Prevalence	0.500	0.202
Naïve Bayes	AUC	0.986	0.988
	Accuracy	0.931	0.965
	Recall	0.882	0.901
	Precision	0.977	0.915
	Specificity	0.980	0.980
	Prevalence	0.500	0.194
Decision Trees	AUC	0.985	0.843
	Accuracy	0.982	0.897
	Recall	0.964	0.842
	Precision	0.999	0.705
	Specificity	0.998	0.911
	Prevalence	0.500	0.202
Logistic Regression	AUC	0.623	0.512
	Accuracy	0.646	0.693
	Recall	0.527	0.441
	Precision	0.691	0.316
	Specificity	0.765	0.757
	Prevalence	0.500	0.202
SCD Classifier	AUC	0.575	0.516
	Accuracy	0.595	0.634
	Recall	0.534	0.481
	Precision	0.608	0.260
	Specificity	0.655	0.671
	Prevalence	0.500	0.194
Extra Tree Classifier	AUC	1.000	0.994

	Accuracy	0.997	0.964
	Recall	0.998	0.970
	Precision	0.996	0.862
	Specificity	0.996	0.963
	Prevalence	0.500	0.194

CONFLICT OF INTEREST

The authors declare that they have no competing interests.

CONTRIBUTION OF INDIVIDUAL AUTHORS

R.K. Kandagatla, V.J. Naidu and P.S. Reddy were involved in design of study analysis and interpretation of data and manuscript revision. P.V. Kavya, M. Joshitha and Ch. Rakesh were involved in conception, design of study, acquisition, analysis of data and writing and drafting of manuscript.

REFERENCES

Anca-Mihaela Vasilica, Vladimir Litvak, Chunyan Cao, Matthew Walker, Umesh Vivekananda. Detection of pathological high-frequency oscillations in refractory epilepsy patients undergoing simultaneous stereo-electroencephalography and magnetoencephalography,Seizure. European Journal of Epilepsy 2023; 107:81-90.

Bajpai R, Yuvaraj A, Prince A, Murugappan M. Abnormal EEG detection using time frequency images and convolutional neural network. Biomedical Signals Based Computer Aided Diagnosis for Neurological Disorders 2022; 133:1–22.

Boashash B, Boubchir.L, Azemi G. A methodology for time-frequency image processing applied to the classification of non- stationary multichannel signals using instantaneous frequency descriptors with application to newborn EEG signals. EURASIP Journal on Advances in Signal Processing 2012, 117(2012): 1–21.

Burgess PW. Riddle is to conundrum as the frontal pole is too..?.Brain 2016; 139 (6): 1627–1630.

Ghosh Dastidar S, Adeli H, Dadmehr N. Mixed band wavelet chaos neural network methodology for epilepsy and epileptic seizure detection. IEEE transactions on biomedical engineering 2007;54(9):1545–1551.

Kumar TS, Kanhangad V, Pachori RB. Classification of seizure and seizure-free EEG signals using local binary patterns. Biomedical Signal Processing and Control 2015; 15:33–40.

Kai Fu, Jianfeng Qu, Chai YI, Dong Y. Classification of seizure based on the time-frequency image of EEG signals using HHT and SVM. Biomedical Signal Processing and Control 2014; 13: 15–22.

Li Y, Wei HL, Billings SA, Liao XF. Time-varying linear and nonlinear parametric model for granger causality analysis. Physical Review E 2012; 85(4): 041906.

Loukas Ilias, Dimitris Askounis. John Psarras, Multimodal detection of epilepsy with deep neural networks. Expert Systems with Applications 2023;213(Part B):119010

Mingkan Shen, Peng Wen, Bo Song, Yan Li. Real-time epilepsy seizure detection based on EEG using tunable-Q wavelet transform and convolutional neural network. Biomedical Signal Processing and Control 2023;82:104566

Olmi B, Frassineti L, Lanata A, Manfredi C. Automatic detection of epileptic seizures in neonatal intensive care units through EEG, ECG and video recordings: A survey. IEEE Access 2021; 9: 138174–138191.

Quanhong Wang, Weizhuang Kong, Jitao Zhong, Zhengyang Shan, Juan Wang, Xiaowei Li, Hong Peng, Bin Hu, ,A hybrid SVM and kernel function-based sparse representation classification for automated epilepsy detection in EEG signals.Neurocomputing 2023.562: 126874

Shu Lih Oh, Jahmunah V, Elizabeth Emma Palmer, Prabal Barua D, Sengul Dogan, Turker Tuncer, Salvador García, Filippo Molinari, Rajendra Acharya U. EpilepsyNet: Novel automated detection of epilepsy using transformer model with EEG signals from 121 patient population.,Computers in Biology and Medicine 2023; 164: 107312

Usman SM, Khalid S, Aslam MH. Epileptic seizures prediction using deep learning techniques. IEEE Access 2020;8: 39998–40007.

Yang Li, Luo ML, Li K. A multiwavelet-based time-varying model identification approach for time–frequency analysis of EEG signals. Neurocomputing 2016; 193:106–114.

Yiying Cai, Kevin Chang, Nuraini Nazeha, Tushar Divakar Gosavi, Jia Yi Shen, Weiwei Hong, Yee-Leng Tan, Nicholas Graves. The cost-effectiveness of a real-time seizure detection application for people with epilepsy, Epilepsy & Behavior 2023;148:109441