IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 6, June 2021

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.542

9940 572 462 | 6381 907 438 | ijircce@gmail.com | www.ijircce.com

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

|| Volume 9, Issue 6, June 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0906118 |

Epileptic Seizure Prediction using EEG Dataset

Pooja Varghat, Sanskruti Gaikwad, Sneha Bagade, Komal Bhosale

Department of Computer Engineering, Bharati Vidyapeeth's College of Engineering for Women, Pune, India

ABSTRACT: In this era, the neurological disease is rapidly increasing worldwide. Also varieties of disease are on rise. Increased demands on health care systems due to neurological disease have become a major concern. Epileptic seizures occur due to disorder in brain functionality which can affect patient's health. Prediction of epileptic seizures before the beginning of the onset is quite useful for preventing the seizure by medication. Machine learning techniques and computational methods are used for predicting epileptic seizures from Electroencephalograms (EEG) signals. However, pre-processing of EEG signals for noise removal and features extraction are two major issues that have an adverse effect on both anticipation time and true positive prediction rate. In that paper a model that provides an methods are pre-processing and feature extraction. Our model predicts epileptic seizures' sufficient time before the onset of seizure starts and provides a better true positive rate. We have applying empirical mode decomposition (EMD) for pre-processing and have extracted time and frequency domain features for training a prediction model.

I.INTRODUCTION

An association between symptoms and diagnosis is majorly encapsulated within the Disease prediction system using EEG signals. The system tends to analyze and predict the disease/disorder based on the symptoms provided by the patient. EPILEPSY is described as a neurological brain disorder identified by the frequent occurrence of symptoms called epileptic seizure due to abnormal brain activities. Seizure's characteristics contain dropping of cognizance or alertness or responsiveness and disruption in movement, sensation or other cognitive functions.

The most affected age group is the people at extremes of age while among young individuals in ages between 10 to 20 years old the disease crests. Epilepsy has a high disease load where millions of people worldwide may have epilepsy and there are about two million new patients recorded every year. Maximum percentage of the epileptic patients could be controlled by the Anti-Epileptic Drugs (AED) while the other 30% are uncontrollable. The electrical recording of the brain activities is Electroencephalogram (EEG) and is considered the most powerful diagnostic and analytical tool of epilepsy. Physicians classify the brain activity of the epileptic patients according to the EEG recordings into four states: preictal state, which is defined by the time period just before the seizure, octal state which is during the seizure occurrence, posticta1 state that is assigned to the period after the seizure took place and finally the interictal state which refers to the period between seizures other than the previously mentioned states. The evaluation of patient-scientific models takes place on the chosen patients by calculating some performance measures accuracy, sensitivity, specificity and warning per hour. This work we develop a device for epileptic seizure detection which alerts the families of epileptic patients to seek emergency services

II.LITERATURE REVIEW

There are many ways to detect the Epileptic Seizure of the any person like Electrocardiography (EEG), Electromyography (EMG), Galvanic Skin Response (GSR), Heart Rate (HR), and Electroencephalography (EEG).Among them EEG is the best, efficient and fast technique use for monitoring the brain. EEG also very safe and risk-free machine. Using EEG signals, there are many things are possible. One of them is Epileptic Seizure detection. There are many systems to detect Epileptic Seizure via EEG but none of them are detect the Epileptic Seizure and also give a recommendation of music or movie according to the Epileptic Seizure

[1] Automated Seizure Prediction" .Authors: U. Rajendra Acharya, Yuki Hagiwara, HojjatAdeli. year: 2018.
In this paper, the authors reviewed challenges faced in seizure prediction, analysis methods, and state-of-the-art techniques used, CNN architecture along with LSTM networks and cloud computing process. Additional research on feature extraction techniques. Further, a more sophisticated machine learning algorithm can be used for a more accurate prediction of seizure.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

|| Volume 9, Issue 6, June 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0906118 |

Finally, it is difficult to achieve an accurate early seizure prediction because of the limited number and length of EEG data. CNN architecture along with LSTM networks.

[2] Epileptic Seizure Detection using deep Learning Approach", VHD Journal of science and technology paper, Author: SirwanTofiqJaafa.Year: 2019.
In this paper, the proposed method consists of two significant folds preprocessing and processing, after preprocessing the preprocessed data are used to train the LSTM network followed by a Softmax function is used to classify the inputs data into normal and seizure data. The preprocessing includes 3 stages, which are normalization of EEG data, applying appropriate filters to select the interesting parts of the data, and data management. The preprocessing includes three stages, which are the normalization of EEG data, applying appropriate filters to select the interesting parts of the data and data management. After preprocessing the preprocessed data are used to train the LSTM network followed by a Softmax function is used to classify the inputs data into normal and seizure data.

[3] Epileptic Seizures Prediction Using Machine Learning Methods"Authors -Syed Muhammad Usman, Muhammad Usman,Simon .Year: 2017.
In this reference paper author have mentioned selection of a suitable classifier, they have performed a classification of the ictal state with rest of states of seizure. In first phase Support Vector Machine is chosen as a classifier due to its superior performance in terms of sensitivity. After selection of classifier, they have applied their model for prediction of epileptic seizures on CHB-MIT dataset that contains EEG signals recordings of several hours .With applying the proposed model on the dataset, on average they have predicted epileptic seizures 23.6 minutes before the start of the onset of a seizure. With the help of this proposed model, epilepsy affected patients will get more time for proper medication required for preventing the seizure before it actually occurs.

[4] Predicting epileptic seizures in scalp EEG based on a variationalbayesiangaussian mixture model of zero-crossing intervals," Authors – A. S. Zandi, R. Tafreshi, M. Javidan, and G. A. Dumont. Year: 2015.
The papers introduce a methodology based on machine learning model for both preprocessing and features extraction. The first stage utilities an algorithm based on scalp EEG zero-crossing intervals in scalp electroencephalogram (EEG). The first phase uses moving-window analysis, the histogram intervals for the current EEG epoch is computed, and the values corresponding to specific bins are selected as an observation. In next phase, the set of observations from the last 5 min is compared with two reference sets of data points (preictal and interictal) through novel measures of similarity and dissimilarity based on a variational Bayesian Gaussian mixture model of the data. A combined index is then computed and compared with a patient-specific threshold, resulting in a cumulative measure which is utilized to form an alarm sequence for each channel. Finally, this channel-based information is used to generate a seizure prediction alarm.

[5] Y. Bengio, I. J. Goodfellow, and A. Courville, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, 2015.
In Decision Trees, Random Forest forms a voting ensemble [5]. Multiple trees are formed on randomly generated input data subsets (Bootstrap Aggregation) and then those trees, the random forest, will all vote on their predicted outcome and a prediction is obtained. Adaptive Boosting is the process of creating multiple unique instances of one form of model prediction so as to effectively improve the model in situations where selected parameters can prove ineffective.

[6] Zheng, Wenming. "Multichannel EEG-based emotion recognition via group sparse canonical correlation analysis."IEEE Transactions on Cognitive and Developmental Systems 9, no. 3 (2017).
They address EEG emotion recognition problem, adopt the common frequency function used to define the EEG signal by dividing the complete EEG frequency band into five parts, and frequency bands, and then extract the frequency band features from each band for GSCCA model learning and emotion recognition. Finally, detailed EEG-based emotion recognition studies based on the SJTU Emotion EEG Dataset (SEED) and experimental findings indicate that the proposed GSCCA system would outperform the state-of-the-art EEG-based approaches to emotion recognition.

[7] Liu, Yong-Jin, Minjing Yu, Guozhen Zhao, Jinjing Song, Yan Ge, and Yuanchun Shi. "Real-time movieinduced discrete emotion recognition from EEG signals." IEEE Transactions on Affective Computing (2017)
Many methods for extraction of features have been studied in paper[7], and the selection of both suitable features and electrode locations is typically based on neuroscientific findings. However, their suitability for emotional recognition was tested using a limited amount of distinct feature sets and on different, typically small data sets. A major drawback here, that

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

|| Volume 9, Issue 6, June 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0906118 |

there is no systematic function comparison. Hence, review feature extraction methods based on 33 studies for EEG emotion recognition. An experiment conducted using machine learning techniques for selecting features on a self-recorded data set to compare those features. Data are provided regarding performance of various methods of selection of features, use of selected types of features and selection of electrode locations. Multivariate-selected features slightly outperform univariate methods. Advanced techniques for extraction of the feature have advantages over commonly used spectral power bands. Results also suggest preference over parietal and centro-parietal lobes for locations. Classification problems are supervised problems of machine learning where the task of predicting a discrete class for a given input was applied. A standard example of handwritten digit recognition where a handwritten digit input picture and a discrete category output. As with all supervised learning problems the training data consists of a series of sample pairs of input-outputs.

[8] Katsigiannis, Stamos, and NaeemRamzan. "DREAMER: A Database for Emotion Recognition Through EEG and ECG Signals from Wireless Lowcost Off-the-Shelf Devices." IEEE Journal of Biomedical and Health Informatics (2017)

All the signals were captured in paper using portable, wearable, wireless, low-cost and off-shelf equipment which has the potential to allow the use of affective computing methods in everyday applications. Through supervised classification experiments using Support Vector Machines (SVMs), a basis for participant-wise emotion recognition was established using EEG and EEG-based features as well as their fusion.

[9] ChinmayiBankar,AdityaBhide ; Anuja Kulkarni ; Chirag GhubeMangeshBedekar. Driving Control(Via music) Using Emotion Analysis Via EEG 2018 IEEE Punecon.27 June 2019.

Proposed a system for driving control using emotion analysis via EEG. They proposed to detection emotion of the driver via EEG then play a music according to emotion. So, driver stay in a positive clam mind. This process is running in a loop. By doing this their aim is reduction of accident. They proposed a Circumplex model for emotion detection. There are two type of EEG Electrode, First is gel electrode and second is dry electrode. They proposed dry electrode. Then according to emotion system will recommend the song for to change mood in the relax mind.

[10] Barjinder Kaur, Dinesh Singh, ParthaPratim Roy. EEG Based Emotion Classification Mechanism in BCIIn: 2018 Procedia Computer Science.Volume 132, 2018, Pages 752-758.

They Proposed a EEG based emotion recognition model. They use SVM machine learning algorithm for emotion detection from EEG data. They successfully detected three emotions (Anger, clam, happy).They use real time data of 10 users. The fractal dimension (FD) feature is fetched by higuchitechnique from EEG signals. Then using SVM with the accuracy of 60% is achieved. They successfully classified happy emotion with the best accuracy. Other two emotion is classified with low accuracy.

III.PROPOSED SYSTEM DESIGN

EEG (electroencephalogram) machine is used for monitoring the brain activity of the person. Many professional doctors uses EEG machine for monitoring the brain activity of the patient. But EEG is also used for may condition like, brain tumor, head injury, seizure disorder, memory problem, brain stroke, dementia, sleep problems etc. Main thing in using EEG is that it is very safe there no risk in using or side effects. There are many types of EEG machines available for different use. Every EEG machine's functionality is different. So, for them, Using EEG (electroencephalogram) to capture the data of the handicapped dump person's brain is easy and relevant. EEG basically used for motoring the brain of the person. Second part of the system is music/movie recommendation system. Here music/movie recommendation system will work as therapy. Physical and anxiety effects are suppressed with the help of music therapy.21 million people in India as suffering from one or the other kind of disability

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

|| Volume 9, Issue 6, June 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0906118 |

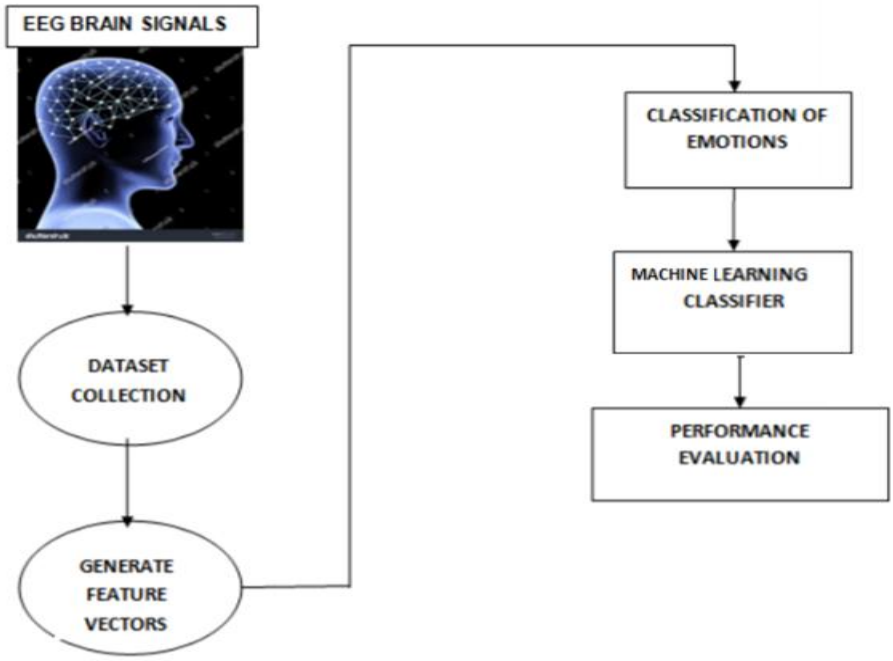


Figure 1 : Proposed system architecture

In proposed research work testing the entire system with supervised learning algorithms, we collect brain data as EEG signals initially. To extract the different features from Seth inputs, and generate the trained model from both machine learning algorithms accordingly. This work is done in two phases, training phase and Test phase training step, use all of the proposed algorithms in each case. The systems first step, then, is to get the persons EEG signals data. Signal pre-processing plays a crucial role in bringing the machine-learning algorithm effective output. The EEG Input signals the values are entered in the machine learning classifier and thee motions are classified by hyper plane using machine learning drawn with the Epileptic Seizure detection identified.

Preprocess

All machine learning algorithms required proper format of dataset. Some noisy data are in the dataset and that noisy data need to be removed by pre-processing. Every data needs to check is there any null values are there or not.

Feature Extraction

To learn the characteristics of expression confiscation from EEG data, machine learning was deployed to eliminate discrimination features related to seizures in EEG. In Feature Extraction the aims to reduce the number of features in a dataset by creating new features from the existing ones. In feature selection the process of reducing the number of input variables when developing a predictive model.The motivation for this was small learning long-term dependence on EEG departments in between each signal and a different EEG signal across same class. Natural language processing (NLP) is the ability of a computer program to understand human language as it is spoken. NLP is a component of artificial intelligence. The correlation patterns are found in the EEG channel, showing the most significant changes over time in terms of feature information retrieval.

Feature selection

In the system, various feature selection approaches are analyzed and hybrid approach for feature selection is proposed.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

|| Volume 9, Issue 6, June 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0906118 |

Classification

After we get the training model, we can feed the testing data into it and get the prediction of classification. The testing stage includes preprocessing of testing data and classification of the testing text. Some classification is required to detect epileptic seizures, an individual can detect epileptic seizures using the stimulation and valence values system.

```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
from sklearn.preprocessing import LabelEncoder
from sklearn.preprocessing import StandardScaler
from sklearn.model_selection import train_test_split, cross_val_score
from sklearn.feature_selection import SelectKBest, RFECV
from sklearn.decomposition import PCA
from sklearn.discriminant_analysis import LinearDiscriminantAnalysis as LDA
from sklearn.linear_model import LogisticRegression
#from xgboost.sklearn import XGBClassifier
from sklearn.metrics import confusion_matrix, accuracy_score, classification_report
import seaborn as sn
#import keras
#from keras.models import Sequential
#from keras.layers import Dense, BatchNormalization, Dropout
#from keras import regularizers
import warnings; warnings.simplefilter('ignore')
%matplotlib inline
dataset = pd.read_csv('C:/Users/HP/Desktop/Project/Epileptic Seizure Recognition.csv')
dataset.head()
```

	Unnamed	X1	X2	X3	X4	X5	X6	X7	X8	X9	...	X170	X171	X172	X173	X174	X175	X176	X177	X178	y
0	X21.V1.791	135	190	229	223	192	125	55	-9	-33	...	-17	-15	-31	-77	-103	-127	-116	-83	-51	4
1	X15.V1.924	386	382	356	331	320	315	307	272	244	...	164	150	146	152	157	156	154	143	129	1
2	X8.V1.1	-32	-39	-47	-37	-32	-36	-57	-73	-85	...	57	64	48	19	-12	-30	-35	-35	-36	5
3	X16.V1.60	-105	-101	-96	-92	-89	-95	-102	-100	-87	...	-82	-81	-80	-77	-85	-77	-72	-69	-65	5
4	X20.V1.54	-9	-65	-98	-102	-78	-48	-16	0	-21	...	4	2	-12	-32	-41	-65	-83	-89	-73	5

5 rows × 180 columns

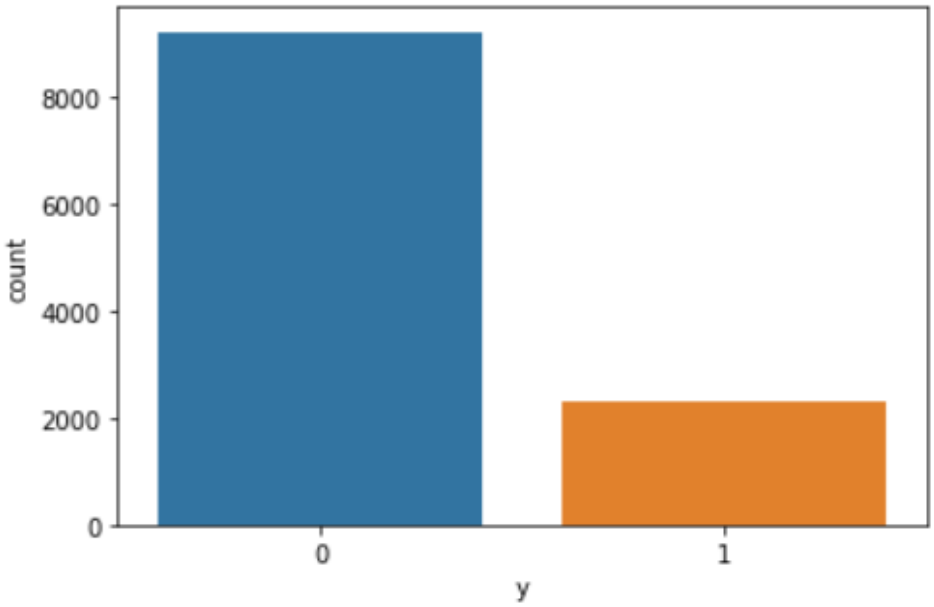
```
cols = dataset.columns
tgt = dataset.y
tgt[tgt>1]=0
ax = sn.countplot(tgt,label="Count")
non_seizure, seizure = tgt.value_counts()
print('The number of trials for the non-seizure class is:', non_seizure)
print('The number of trials for the seizure class is:', seizure)
```

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

|| Volume 9, Issue 6, June 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0906118 |

```
The number of trials for the non-seizure class is: 9200
The number of trials for the seizure class is: 2300
```



As we can see, there are 178 EEG features and 5 possible classes. The main goal of the dataset it's to be able to correctly identify epileptic seizures from EEG data, so a binary classification between classes of label 1 and the rest (2, 3, 4, 5). In order to train our model, let's define our independent variables (X) and our dependent variable (y).

```
(11500, 178)
```

```
y = dataset.iloc[:,179].values
y
```

```
array([0, 1, 0, ..., 0, 0, 0], dtype=int64)
```

To make this a binary problem, let's turn the non-seizure classes 0 while maintaining the seizure as 1.

```
y[y>1]=0
y
```

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

|| Volume 9, Issue 6, June 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0906118 |

```
array([0, 1, 0, ..., 0, 0, 0], dtype=int64)
```

Let's split our data between training and test. We later will employ 10-fold cross validation to ensure that our model correctly generalize to unseen data.

Train and test splits

```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.2)
```

Applying scaling to our independent variables to make easier for our model to converge.

Scaling the data

```
sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)
```

```
X_train[0,0:30]
```

```
array([0.03898682, 0.01909605, 0.02902387, 0.17218818, 0.53118163,
       1.09703726, 1.65633723, 2.10065872, 2.20983457, 2.12669069,
       1.88204962, 1.57188419, 1.27781999, 1.09512026, 1.04632091,
       1.07301766, 1.09719483, 1.11313783, 1.0815289 , 0.98510555,
       0.91357687, 0.8815188 , 0.84517751, 0.87375587, 0.90231337,
       0.92034803, 0.92127195, 0.89649028, 0.91147618, 0.95474192])
```

```
y_train[0:30]
```

```
array([0, 0, 1, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1, 0, 0,
       0, 0, 0, 1, 0, 0, 0, 0], dtype=int64)
```

Let's investigate how accurate the simplest model can achieve by applying a Logistic Regression to our data.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

|| Volume 9, Issue 6, June 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0906118 |

Simple Classifier

```
accuracies = cross_val_score(estimator = classifier,
                             X = X_train,
                             y = y_train,
                             cv = 10,
```

```
y_pred = classifier.predict(X_test)
```

```
LogisticRegression(C=1.0, class_weight=None, dual=False, fit_intercept=True,
                   intercept_scaling=1, l1_ratio=None, max_iter=100,
                   multi_class='warn', n_jobs=None, penalty='l2',
                   random_state=None, solver='warn', tol=0.0001, verbose=0,
                   warm_start=False)
```

	precision	recall	f1-score	support
Non-seizure	0.83	1.00	0.91	1865
Seizure	0.98	0.11	0.20	435
accuracy			0.83	2300
macro avg	0.90	0.55	0.55	2300
weighted avg	0.86	0.83	0.77	2300

```
classifier = LogisticRegression()
classifier.fit(X_train, y_train)
```

```
0.8142380839296832
```

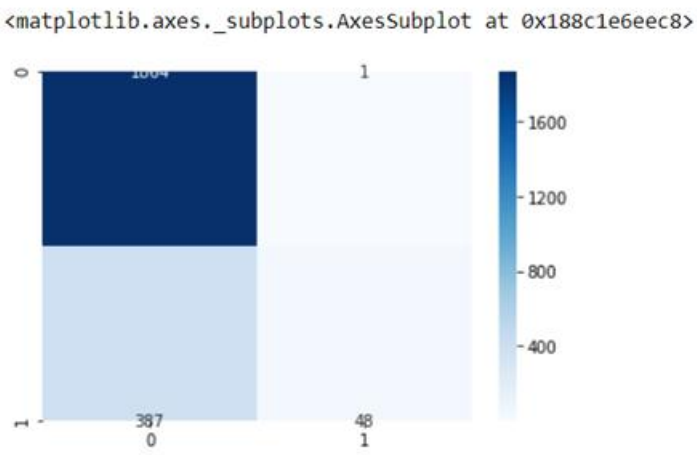
```
cm = confusion_matrix(y_test, y_pred)

print(classification_report(y_test, y_pred, target_names=['Non-seizure', 'Seizure']))
```

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542 |

|| Volume 9, Issue 6, June 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0906118 |



IV.CONCLUSION AND FUTURE WORK

This project aims to predict the disease on the basis of the symptoms. In this project the system takes symptoms from the user as input and produces output i.e. predict disease and recommendation of herbs' product to the user. We are going to provide more information about the herbs' product. This system will also help the doctors analyze the pattern of presence of diseases in the society and can also be used for marketing of herb's product. In future we can add large amount of Data and we can use Map Reduce to handle the large amount of Data.

REFERENCES

[1] Automated Seizure Prediction" .Authors: U. Rajendra Acharya, Yuki Hagiwara, Hojjat Adeli. year: 2018.
[2] Epileptic Seizure Detection using deep Learning Approach", VHD Journal of science and technology paper, Author: Sirwan Tofiq Jaafa.Year :2019.
[3] Epileptic Seizures Prediction Using Machine Learning Methods"Authors -Syed Muhammad Usman, Muhammad Usman,Simon .Year: 2017.
[4] Predicting epileptic seizures in scalp EEG based on a variational bayesian gaussian mixture model of zero-crossing intervals," Authors – A. S. Zandi, R. Tafreshi, M. Javidan, and G. A. Dumont. Year:2015
[5] Y. Bengio, I. J. Goodfellow, and A. Courville, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, 2015.
[6] Zheng, Wenming. "Multichannel EEG-based emotion recognition via group sparse canonical correlation analysis."IEEE Transactions on Cognitive and Developmental Systems 9, no. 3 (2017).
[7] Liu, Yong-Jin, Minjing Yu, Guozhen Zhao, Jinjing Song, Yan Ge, and Yuanchun Shi. "Real-time movieinduced discrete emotion recognition from EEG signals." IEEE Transactions on Affective Computing (2017).
[8] Katsigiannis, Stamos, and NaeemRamzan. "DREAMER: A Database for Emotion Recognition Through EEG and ECG Signals from Wireless Lowcost Off-the-Shelf Devices." IEEE Journal of Biomedical and Health Informatics (2017).
[9] Chinmayi Bankar,Aditya Bhide ; Anuja Kulkarni ; Chirag Ghube Mangesh Bedekar.Driving Control(Via music) Using Emotion Analysis Via EEG 2018 IEEEPunecon.27 June 2019.
[10] Barjinder Kaur, Dinesh Singh, Partha Pratim Roy. EEG Based Emotion ClassificationMechanism in BCIIn:2018 Procedia Computer Science.Volume 132, 2018, Pages 752-758.

INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 7.542

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

Scan to save the contact details