MACHINE LEARNING FRAMEWORK FOR IDENTIFICATION OF ABNORMAL EEG SIGNAL

A Mini Project Report Submitted in Partial Fulfillment of the Requirement for the Award of the Degree of

BACHELOR OF TECHNOLOGY

in

ELECTRONICS AND COMMUNICATION ENGINEERING

by

JELLA BHASKAR	**19PA1A0465**
MANUSU MADHU BABU	**19PA1A04A3**
MANDA ROHIT SATYA SIDDU MANIKANTA	**19PA1A04A0**
MUDUNURI BALAJI SUNIL VARMA	**19PA1A04A7**
KODAMANCHILI MAHIMA	**19PA1A0477**

Under the Esteemed Guidance of

Dr. A. PRABHAKARA RAO

Professor

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

VISHNU INSTITUTE OF TECHNOLOGY (Autonomous)

(Accredited by NBA, NAAC, Approved by AICTE & Affiliated to JNTU Kakinada)

Vishnupur, Bhimavaram - 534202.

2021 - 2022

VISHNU INSTITUTE OF TECHNOLOGY (AUTONOMOUS)

(Accredited by NBA, NAAC, Approved by AICTE & Affiliated to JNTU Kakinada)
Vishnupur, Bhimavaram-534202

DEPARTMENT OF ELECTRONICS AND COMMUNATION ENGINEERING

CERTIFICATE

This is to certify that the Mini Project entitled **"Machine learning Framework for Identification of Abnormal EEG Signal"** is being submitted by **JELLA BHASKAR (19PA1A0465), MANASU MADHU BABU (19PA1A04A3), MANDA ROHIT SATYA SIDDU MANIKANTA (19PA1A04A0), MUDUNURI BALAJI SUNIL VARMA (19PA1A04A7), KODAMANCHILI MAHIMA (19PA1A0477)** in partial fulfillment for the award of the degree of **Bachelor of Technology in Electronics and Communication Engineering** is a record of the bonafide work carried out by them under my guidance and supervision during academic year 2021–2022 and it has been found worthy of acceptance according to the requirements of the university.

Dr. A. Prabhakara Rao
Professor
Project Guide

Dr. N. Padmavathy, Ph. D (IIT- KGP)
Professor & Head

External Examiner

ABSTRACT

Epilepsy is a chronic noncommunicable disease of the brain that affects people of all ages. It may lead to seizures, loss of awareness, unusual sensations and behavior. Around 50 million people worldwide have epilepsy, making it one of the most common neurological diseases globally according to WHO statistics in 2021. It is estimated that up to 70% of people living with epilepsy could live seizure- free if properly diagnosed and treated. Electroencephalograms (EEG) are universally used to this chronic noncommunicable disease. However, the evaluation of specific variety of abnormality utilizing the EEG signal is an instinctive event and may differ from radiologist-to- radiologist. Visual examination of the EEG signal by observing a change in frequency or amplitude in long-duration signals is a difficult task for the radiologists. It may give rise to inaccurate categorization. Determination of epileptic seizure is essential task for the treatment of epilepsy. This project proposes a machine learning Framework that can detect the abnormality in the EEG signal automatically to assist the radiologists in their diagnosis. The ML Framework consist of 7 classifiers i.e., KNN, SVM, Random Forest, Logistic Regression, Decision Tree, AdaBoost, Bagging. Wavelets are well localized both in time and frequency domains, so they are used to extract features from bonn dataset by multi-level (4) wavelet decomposition & the extracted features fed to ML Framework for the detection of abnormal EEGs. Out of 7 classifiers in the Framework bagging classifier perform well for the detection of abnormal EEGs with 70% accuracy & 76% AUC

Keywords: Epilepsy, Seizure, Electroencephalograms (EEG), Radiologist, Wavelets, KNN, SVM, Random Forest, Logistic Regression, Decision Tree, AdaBoost, Bagging, Framework, Accuracy, AUC.

CONTENTS

CHAPTER I

INTRODUCTION

INTRODUCTION

Epilepsy is one of the most common neurological diseases globally according to WHO statistics in 2021. Nearly 50 million people get affected due to it according to WHO statistics up to 2021.It causes seizure, Stiff muscles etc. Prolonged seizures are clearly capable of injuring the brain. Isolated, brief seizures are likely to cause negative changes in brain function and possibly loss of specific brain cells. It is estimated that up to 70% of people living with epilepsy could live seizure- free if properly diagnosed and treated [21]

1.1 EPILEPSY

The word epilepsy originates from the Latin and Greek word ‘epilepsia’ which means ‘seizure’ or ‘to seize upon’. It is a serious neurological disorder with unique characteristics, tending of recurrent seizures. The context of epilepsy, found in the Babylonian text on medicine, was written over 3000 years ago. This disease is not limited to human beings, but extends to cover all species of mammals such as dogs, cats and rats. However, the word epilepsy does not give any type of clues about the cause or severity of the seizures; it is unremarkable and uniformly distributed around the world.

Several theories about the cause are already available. The main cause is electrical activity disturbance inside a brain, which could be originated by several reasons such as malformations, shortage of oxygen during childbirth, and low sugar level in blood. Globally, epilepsy affects approximately 50 million people, with 100 million being affected at least once in their lifetime. Overall, it accounts for 1% of the world’s burden of diseases, and the prevalence rate is reported at 0.5–1%. The main symptom of epilepsy is to experience more than one seizure by a patient. It causes a sudden breakdown or unusual activity in the brain that impulses an involuntary alteration in a patient’s behavior, sensation, and loss of momentary consciousness. Typically, seizures last from seconds to a few minute(s), and can happen at any time without any aura. This leads to serious injuries including fractures, burns, and sometimes death.

Based on the symptoms, seizures are categorized by neuro-experts into two main categories—partial and generalized. Partial seizure, also called ‘focal seizure’, causes only a section of the cerebral hemisphere to be affected. There are two types of Partial seizure: simple-partial and complex-partial. In the simple-partial, a patient does not lose consciousness but cannot communicate properly. In the complex-partial, a person gets confused about the surroundings and starts behaving abnormally like chewing and mumbling; this is known as ‘focal impaired awareness seizure’. On the contrary, in the generalized seizures, all regions of the brain suffer and entire brain networks get affected quickly. Generalized seizures are of many types, but they are broadly divided into two categories: convulsive and non-convulsive.[1] The below figure 1 shows the part of the brain effected by various types of seizures.

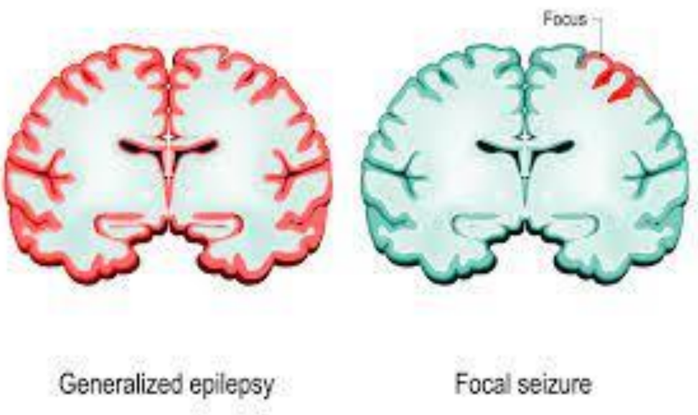


Fig 1: Types of Seizures [8]

1. Focal Seizures/Partial Seizures: Focal seizures are the epileptic seizures in which the abnormal brain activity begin in one specific part of the brain.[1]

They can be segregated into two sub-categories:

(i) Focal seizures with impaired awareness/ Complex partial seizures: These types of seizures involve loss of awareness or consciousness. During the focal seizure with impaired awareness a patient might be unable to do repetitive action such as walking in circles or rubbing hands and might not respond to normal environmental stimuli.[1]

(ii) Focal seizures without loss of consciousness/Simple Partial Seizures: These types of seizures are characterized by retained consciousness or awareness, meaning that consciousness is not altered. Although such kind of seizures may lead to altering the sensory inhibitions or emotions. Such type of seizures may also cause jerking movement of arms and legs. [1]

2. Generalized seizures: Generalized seizures are those type of seizures in which the abnormal brain activity arises from all areas of the brain. The generalized seizures are further sub-characterized into 6 types:

(i) Absence seizures: Absence seizures are generalized seizures also known as petit mal seizures usually occur in children and may by staring blankly in open space or small body movements such as lip-smacking or eye blinking.[1]

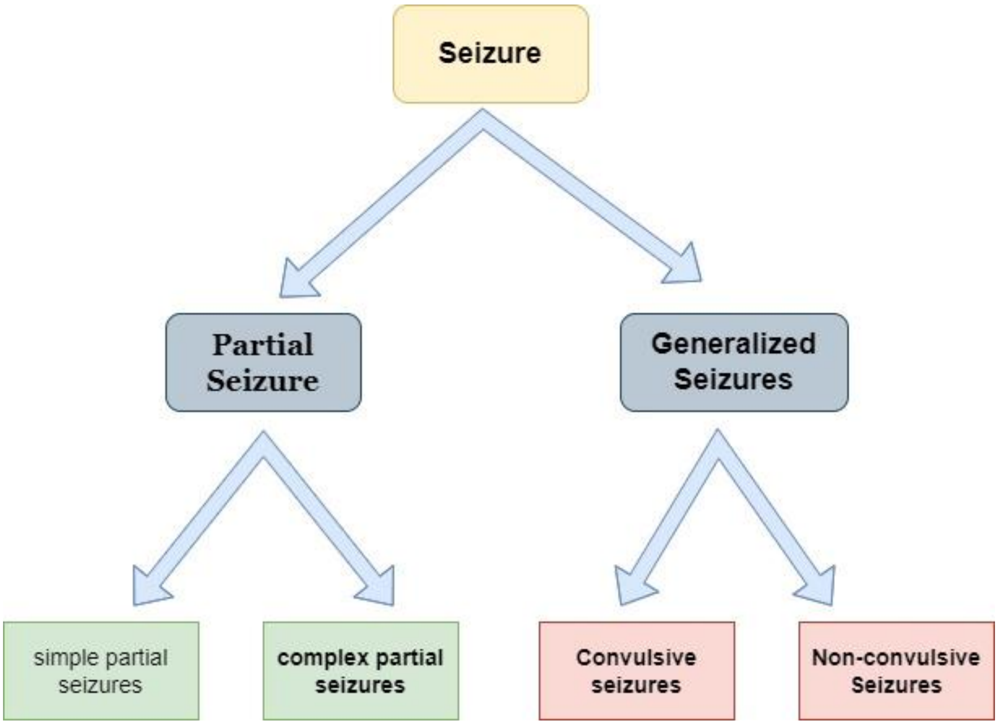


Fig 2: Classification of Seizures

(ii) Atonic seizure: are also known as drop seizures are characterized by the sudden loss of muscle control, which may lead to collapsing.[1]

(iii) Myoclonic seizure: are often seen as sudden or brief quick jerking movements of the arms and legs.

(iv) Clonic seizure: are characterized by repeated jerking of limbs. The neck, face and other body parts may also get affected by these seizures.[1]

(v) Tonic seizure: may lead to muscle stiffening, especially in the back, neck, and limbs, causing the individual to fall down.[1]

(vi) Tonic-clonic seizure/ Convulsive seizures: also known as grand mal seizures. This type of seizures may lead to abrupt jerking and stiffening of muscles. It may also cause loss of consciousness, bladder control, and biting of the tongue [1]

(v) Non-convulsive: Nonconvulsive status epilepticus (NCSE) refers to a prolonged seizure that manifests primarily as altered mental status as opposed to the dramatic convulsions seen in generalized tonic-clonic status epilepticus. There are 2 main types of NCSE, each of which has a different presentation, cause, and expected outcome.

1.2 ELECTROENCEPHALOGRAPHY

Electroencephalography (EEG) is an electrophysiological monitoring method to record electrical activity of the brain. It is typically non-invasive, with the electrodes placed along the scalp, although invasive electrodes are sometimes used in specific applications. EEG measures voltage fluctuations resulting from ionic current within the neurons of the brain. In clinical contexts, EEG refers to the recording of the brain's spontaneous electrical activity over a period of time, as recorded from multiple electrodes placed on the scalp. Diagnostic applications generally focus on the spectral content of EEG, that is, the type of neural oscillations (popularly called "brain waves") that can be observed in EEG signals. The test setup for EEG & recording of various EEG signals is shown in the figure 3 below.

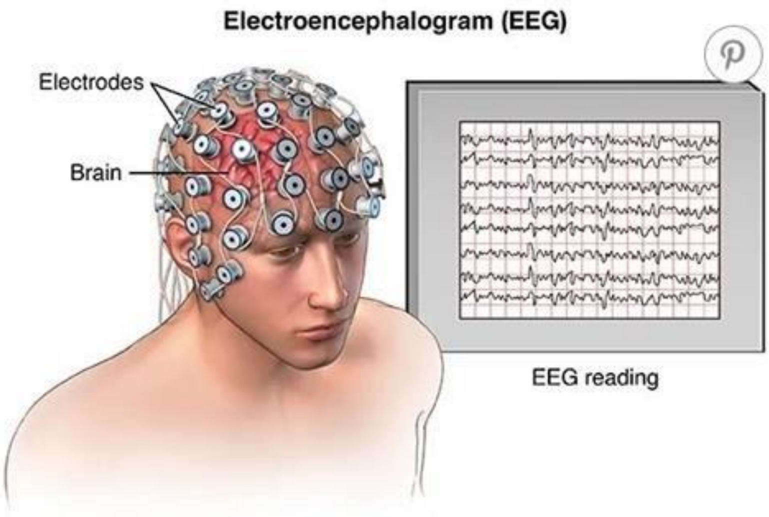


Fig 3: Electroencephalogram [9]

EEG is most often used to diagnose epilepsy, which causes abnormalities in EEG readings. It is also used to diagnose sleep disorders, coma, encephalopathy, and brain death. EEG can be used as a first-line method of diagnosis for tumors, stroke and other focal brain disorders, but this use has decreased with the advent of high-resolution anatomical imaging techniques such as magnetic resonance imaging (MRI) and computed tomography (CT). Despite limited spatial resolution, EEG continues to be a valuable tool for research and diagnosis, especially when millisecond-range temporal resolution (not possible with CT or MRI) is required.[3] The figure 4 shows the EEG cap fix to the head of the patient.

Fig 4: An EEG setup [10]

A routine clinical EEG recording typically lasts 20–30 minutes (plus preparation time) and usually involves recording from scalp electrodes. Routine EEG is typically used in the following clinical circumstances:

1. To distinguish epileptic seizures from other types of spells, such as psychogenic non- epileptic seizures, syncope, sub-cortical movement disorders and migraine variants.
2. To differentiate "organic" encephalopathy or delirium from primary psychiatric syndromes such as catatonia.
3. To serve as an adjunct test of brain death.
4. To prognosticate, in certain instances, in patients with coma.
5. To determine whether to wean anti-epileptic medications.

Additionally, EEG may be used to monitor certain procedures:

1. To monitor the depth of anesthesia.
2. As an indirect indicator of cerebral perfusion in carotid endarterectomy.
3. To monitor am barbital effect during the Wada test.

EEG can also be used in intensive care units for brain function monitoring:

1. To monitor for non-convulsive seizures/non-convulsive status epileptics
2. To monitor the effect of sedative/anesthesia in patients in medically induced coma (for treatment of refractory seizures or increased intracranial pressure)
3. To monitor for secondary brain damage in conditions such as subarachnoid hemorrhage (currently a research method)

1.2.1 EEG Signal Classification

EEG signals are basically classified into 4 types: Delta, Theta, Alpha and Beta. Recently, two more types have been included – Mu and Gamma. Table 1.1 shows the classification, frequency ranges and importance of different types of EEG signals.

1.2.2 Advantages

Several other methods to study brain function exist, including functional magnetic resonance imaging (fMRI), positron emission tomography, magneto encephalography (MEG), Nuclear magnetic resonance spectroscopy, Electrocorticography, Single-photon emission computed tomography, Near-infrared spectroscopy (NIRS), and Event related optical signal (EROS).

Table 1.1: Classification of EEG signals [2]

Type	Frequency (Hz)	Location	Normally
Delta	Up to 4	Frontally in adults, posterior in children, high amplitude waves	• adults slow sleep wave • in babies • has been found during some continuous attention tasks
Theta	4 - 8	Found in locations not related to task at hand	• young children • drowsiness or arousal in older children and adults • idling • associated with inhibition of elicited responses (has been found to spike in situations where a person is actively trying to repress a response or action).
Alpha	8 - 13	Posterior regions of head, both sides, higher in amplitude on non-dominant side. Central sites (C3-C4) at rest.	• relaxed/reflecting • closing the eyes • Also associated with inhibition control, seemingly with the purpose timing inhibitory activity in different locations across the brain.

Beta	>13 - 30	Both sides, symmetrical distribution, most evident frontally, low amplitude waves	• alert/working • active, busy or anxious thinking, active concentration
Gama	30 - 100+	Somatosensory cortex	• Displays during cross-modal sensory processing (perception that combines two different senses, such as sound and sight) • Also is shown during short term memory matching of recognized objects, sounds, or tactile sensations
Mu	8 - 13	Sensorimotor cortex	• Shows rest state motor neurons.

Despite the relatively poor spatial sensitivity of EEG, it possesses multiple advantages over some of these techniques:

1. Hardware costs are significantly lower than those of most other techniques
2. EEG prevents limited availability of technologists to provide immediate care in high traffic hospitals.
3. EEG sensors can be used in more places than fMRI, SPECT, PET, MRS, or MEG, as these techniques require bulky and immobile equipment. For example, MEG requires equipment consisting of liquid helium-cooled detectors that can be used only in magnetically shielded rooms, altogether costing upwards of several million dollars; and fMRI requires the use of a 1-ton magnet in, again, a shielded room.
4. EEG has very high temporal resolution, on the order of milliseconds rather than seconds. EEG is commonly recorded at sampling rates between 250 and 2000 Hz in clinical and research settings, but modern EEG data collection systems are capable of recording at sampling rates above 20,000 Hz if desired. MEG and EROS are the only other noninvasive cognitive neuroscience techniques that acquire data at this level of temporal resolution.
5. EEG is relatively tolerant of subject movement, unlike most other neuroimaging techniques. There even exist methods for minimizing, and even eliminating movement artifacts in EEG data.
6. EEG is silent, which allows for better study of the responses to auditory stimuli.
7. EEG does not aggravate claustrophobia, unlike fMRI, PET, MRS, SPECT, and sometimes MEG.
8. EEG does not involve exposure to high-intensity (>1 tesla) magnetic fields, as in some of the other techniques, especially MRI and MRS. These can cause a variety of undesirable issues with

the data, and also prohibit use of these techniques with participants that have metal implants in their body, such as metal-containing pacemakers.

9. EEG does not involve exposure to radioligands, unlike positron emission tomography. Event Related Potential (ERP) studies can be conducted with relatively simple paradigms, compared with IE block- design fMRI studies.
10. Extremely un-invasive, unlike Electro-corticography, which actually requires electrodes to be placed on the surface of the brain.

EEG also has some characteristics that compare favorably with behavioral testing:

1. EEG can detect covert processing (i.e., processing that does not require a response).
2. EEG can be used in subjects who are incapable of making a motor response.
3. Some ERP components can be detected even when the subject is not attending to the stimuli
4. Unlike other means of studying reaction time, ERPs can elucidate stages of processing.
5. EEG is a powerful tool for tracking brain changes during different phases of life. EEG sleep analysis can indicate significant aspects of the timing of brain development, including evaluating adolescent brain maturation. Brain activity can also be monitored by Computed tomography (CT scan).
6. In EEG there is a better understanding of what signal is measured as compared to other research techniques, i.e., the BOLD response in MRI.

1.2.3 Disadvantages

1. Low spatial resolution on the scalp. fMRI, for example, can directly display areas of the brain that are active, while EEG requires intense interpretation just to hypothesize what areas are activated by a particular response.
2. EEG poorly measures neural activity that occurs below the upper layers of the brain (the cortex).
3. Often takes a long time to connect a subject to EEG, as it requires precise placement of dozens of electrodes around the head and the use of various gels, saline solutions, and/or pastes to keep them in place. While the length of time differs dependent on the specific EEG device used, as a general rule it takes considerably less time to prepare a subject for MEG, fMRI, MRS, and SPECT.
4. Signal-to-noise ratio is poor, so sophisticated data analysis and relatively large numbers of subjects are needed to extract useful information from EEG.

1.3 WAVELET

Fourier transforms deal with signals that don't have compact support and can be thought of as a translation between functions of the same type: it's a unitary map on an inner product space. Fourier series don't have this property which makes them so much harder to study in full detail.so Fast Fourier Transform (FFT) are used. But a disadvantage associated with the FFT is the restricted range of waveform data that can be transformed and the need to apply a window weighting function (to be defined) to the waveform to compensate for spectral leakage (also to be defined). This kind of signal decomposition may not serve all applications well, for example Electroencephalogram (EEG) where signals have short intervals of characteristic oscillation. An alternative approach is the **Wavelet Transform**, which **decomposes a function into a set of wavelets**. A **Wavelet** is a **wave-like oscillation that is localized in time**.

By using wavelet Transform we can overcome the problem with Short Time Fourier Transform in wavelet transform we used different window size for different frequency components. Low scale (small window size or small-time scale) is used for higher frequencies and higher scale (Large Window size or large time scale) is used for low frequencies [7]

The main advantage of wavelet transforms is

1. Wavelet transforms has multi resolution property
2. Better Spectral localization property

What is Multi resolution property?

Multi resolution property means different frequency components present in any signal are resolved at different scale (different time scale or different window size) Scale is inversely proportional to the frequency means small scale is used for higher frequencies whereas large scale is used for small frequencies small scale (Small window Size or Small-time scale) is used for the analysis of higher frequency components. Means wavelet transform provide higher time resolution for high frequency means if any signal contains a very high frequency component, then with the help of wavelet transform, we can know that at which exact time interval (very small-time interval) these frequency components exists Whereas the large scale (large window size or large time scale) is used for the analysis of small frequency components. Wavelet transform provide good time resolution for higher frequencies whereas for small frequencies time resolution is not so good but if we study real word signals then we find that generally higher frequencies are occurs only for very short time interval whereas small frequency components are presents for long time interval. So, wavelet transform proves its suitability for real time signals.

What is SCALE?

Scale is inversely proportional to the frequency of any signal. Large scale is used for the analysis of small frequency components presents in any signal whereas small scale is used for the analysis of high frequency components of any signal [7].

What is Spectral Localization?

Spectral localization property means that wavelet transform tells us that what frequency components are present in any given signal and at time axis where these frequency components are presents The Wavelet transform has multi resolution capability it means it resolve the different frequency components of the signal at different scales. Scale is inversely proportional to the frequency of the signal it means the high frequency signals are resolved at low scale whereas low frequency signals are resolved at high scale because by choosing a single scale we cannot resolve the all-frequency components present in any signal or in other words if we want to capture every detail of the signal then we need to resolve the different frequency components present in signal at different scale.

To understand the concept of scale, take the following example we can draw the map of globe which shows only where Sea (water) exist and where earth exists. If we can resolve more this map then we will be able to see the boundary of each nation. If we can more resolve this map then we will be able to see the boundaries of states of each nation. If we can further increase resolution then we will be able to see the map of major cities of countries and if we further increased resolution then we will be able to see the different sector, blocks or different area of the city. So, as we are increasing resolution then we are able to know each detail of map or each detail of locations. Or in the other word we can say that by choosing a single resolution level we cannot capture every detail in the map for capturing the different detail of map we need to resolve the map at different levels this is called multi resolution.

Again, take another example suppose we are taking the photograph of city from the peak of any mountain then the whole city will look like a dense population area or look like that the group of several houses again if we zoom our camera then we can focus any sector of city we can also watch the roads, garden, electric cables etc. Again, if we can further zoom our camera then we can focus at any particular home and can watch that particular home how many windows are there in any particular home or how many doors that particular home has again we can focus at the roof of any particular home and even focus at the face of any person who is standing at the roof. We can also zoom at the name-plate in front of any house and can read the name of people. So, by zooming our camera at different- different levels we can able to capturing the different level of details by choosing a single zooming level we can’t capture every detail of any picture The same thing with wavelet transforms in wavelet transform we resolve each frequency components at different -different scales.

The mathematical expression for wavelet transform is

Continuous Wavelet Transform (CWT)

$$T(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t)\Psi^{*} \frac{(t-b)}{a} dt \ldots\ldots\ldots 1.1$$

Discrete Wavelet Transform (DWT)

$$T_{m,n} = \int_{-\infty}^{+\infty} x(t)\psi_{m,n}(t)\, dt \ldots\ldots\ldots 1.2$$

Where:

a = scale of the wavelet

b = location of the wavelet/translation parameter

$\psi(t)$ = transforming wavelet/mother wavelet

CHAPTER 2

LITERATURE SURVEY

LITERATURE SURVEY

In the literature, many studies have been conducted to create a system that accurately detects the abnormal EEGs in human beings as correct classification of such EEGs may lead to the detection of epilepsy, sleep disorders, etc. These Studies not just focus on detection of seizures but also on the prediction of seizures. Due to the non-linear and non-stationary nature of EEG signals, linear analysis methods are not very reliable and may not lead to accurate results. Hence, various machine learning (ML) or deep learning techniques with non-linear analysis methods are used in different studies. In the studies using ML, different features are extracted. Few common ones are, wavelet transform, Hilbert-Huang transform, Eigenvalue decomposition, higher-order spectra and cumulant features. These studies may use single channel or multi-channel signals. In ML, various classifiers like k nearest neighbor (KNN), support vector machine (SVM), random forest (RF), bagged trees, etc. are used to classify the signals on the basis of their signature or extracted features [6]. Further, component analysis can also be used in the classification of EEGs. Lopez et al. [4] used principal component analysis (PCA) with KNN and RF classifier and obtained 58.2% accuracy with KNN classifier and 68.3% with RF classifier.

ML and signal processing usually begins with normalization or standardizing of signals following the pre-processing of signals to remove various artifacts and noise. After the pre-processing stage, the features are extracted. Then, the extracted features are fed into the classifiers for classification and the classification performance is observed. If satisfactory performance is obtained from a model, then it is used for further testing using a different independent data set. Recently various deep learning-based techniques, which do not require feature extraction and selection, are also being employed in studies these days. Acharya et al. [2] used 1D convolutional neural network (CNN) using single-channel signals of duration 60 s and obtained an error rate of 20.6% in classifying the abnormal EEG signals using the same database (i.e., Temple university dataset). Lopez [1] proposed a system that used both 2D CNN and ML on four-channel signals, which resulted in 21.2% error rate for the classification of abnormal patterns in EEG signals. Acharya et al. [20] (13 layer) proposed a CNN-based approach for automated identification of seizure and non-seizure EEG signals. The study resulted in the classification accuracy of 88.67% using 300 signals from 5 patients. Oh et al. [5] also presented a method for the detection of Parkinson’s disease (PD) using a 13-layer CNN model. Their study used EEG signals from 20 healthy and 20 PD patients and obtained an accuracy of 88.25%

CHAPTER 3

METHODOLOGY

METHODOLOGY

The process involved in this project work is depicted in the flowchart shown in figure 6′ below

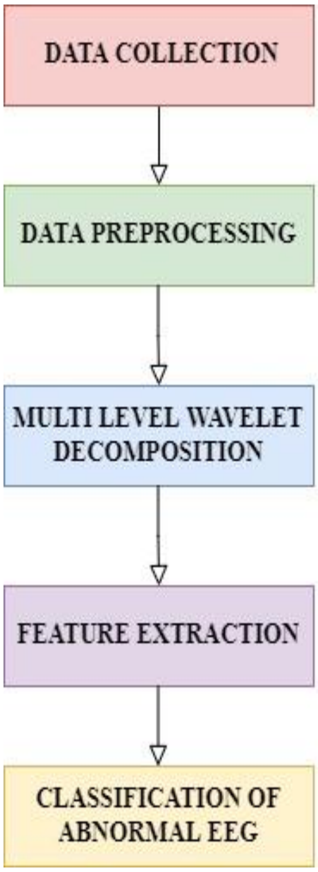


Fig 6′ : Flow work of proposed model

3.1 DATA COLLECTION

The epileptic seizure EEG data is collected from Bonn dataset which is a open source data.

The dataset consists of 5 different folders, each with 100 files, with each file representing a single subject/person. Each file is a recording of brain activity for 23.6 seconds. The corresponding time-series is sampled into 4097 data points. Each data point is the value of the EEG recording at a different point in time. So, we have total 500 individuals with each has 4097 data points for 23.5 seconds. We divided and shuffled every 4097 data points into 23 chunks, each chunk contains 178 data points for 1 second, and each data point is the value of the EEG recording at a different point in time. So now we have 23 x 500 = 11500 pieces of information(row), each information contains 178 data points for 1 second(column), the last column represents the label y {1,2,3,4,5}. The response variable is y in column 179, the Explanatory variables X1, X2, ..., X178 y contains the category of the 178-dimensional input vector. Specifically, y in {1, 2, 3, 4, 5}:

5 - eyes open, means when they were recording the EEG signal of the brain the patient had their eyes open

4 - eyes closed, means when they were recording the EEG signal the patient had their eyes closed

3 - Yes, they identify where the region of the tumor was in the brain and recording the EEG activity from the healthy brain area

2 - They recorder the EEG from the area where the tumor was located

1 - Recording of seizure activity.

Hence 1,2,3 represent unhealthy person and 4,5 are healthy person.

3.2 DATA PREPROCESSING

In this ML Framework we consider healthy record (4,5) and unhealthy record (1,2,3) which represent seizure activity. Initially the sampling frequency of the dataset is 178.3Hz .so we resample the two records to 128 Hz frequency. In the resampling process two filters are used they are bandpass filter & notch filter.

An ideal bandpass filter would have a completely flat passband: all frequencies within the passband would be passed to the output without amplification or attenuation, and would completely attenuate all frequencies outside the passband. In practice, no bandpass filter is ideal. As a =result, we have filter roll-off. To eliminate this filter roll-off we use notch filter as it blocks a specific band of frequencies and allow all frequencies outside the band.

A bandpass filter is an electronic device or circuit that allows signals between two specific frequencies to pass, but that discriminates against signals at other frequencies. Some bandpass filters require an external source of power and employ active components such as transistors and integrated circuits; these are known as active bandpass filters. Other bandpass filters use no external source of power and consist only of passive components such as capacitors and inductors; these are called passive bandpass filters.

What is notch filter?

A notch filter (also known as a band stop filter or reject filter) is defined as a device that rejects or blocks the transmission of frequencies within a specific frequency range and allows frequencies outside that range. Notch filters eliminate transmission of a narrow band of frequencies and allow transmission of all the frequencies above and below this band. As it eliminates frequencies hence, it is also called a band elimination filter.

3.3 MULTILEVEL WAVELET DECOMPOSITION

Using wavelet decomposition technic, it is possible to decompose a signal into a series of orthogonal wavelets. A multiresolution representation of provides a simple hierarchical framework to analyze the signal at different resolution level. This is similar to the notion of decomposing a signal in terms of Fourier_transform components or Walsh or Haar functions. Orthogonality represents a unique and complete representation of a signal. Based on Mallat theory, multiresolution representation of a signal is effective in analyzing information content of a signal in different detail level (Mallat, 1989). This operator is capable of approximating a signal at a given resolution.

In the process of wavelet transform of a signal (S) is first decompose into Approximate Coefficients and Detailed Coefficients by simply passing the signal through low pass filter and high pass filter respectively. The output of low pass filter is called Approximate [A1] (Low frequency components) coefficient of the signal The output of High pass filter is called Detailed [D1] (High frequency components) Coefficients Of the signal.

This Approximate coefficient [A1] again passed through a low pass and high pass filter And again decompose the signal into Approximate [A2] and Detailed Coefficients [D2] Further Approximate Components [A2] can be decomposed into approximate coefficients [A3] And Detailed Coefficients [D3] The number of Decomposition levels depends on the length of signal and our requirements.

S =A1+D1 [First level Wavelet Decomposition] …... (1)

A1=A2+D2 [Second level Wavelet Decomposition] … (2)

A2=A3+D3 [Third level Wavelet Decomposition] …. (3)

S=A3+D3+D2+D1 ………………………………… (4)

The block diagram shown in figure 6 represents the multilevel decomposition.

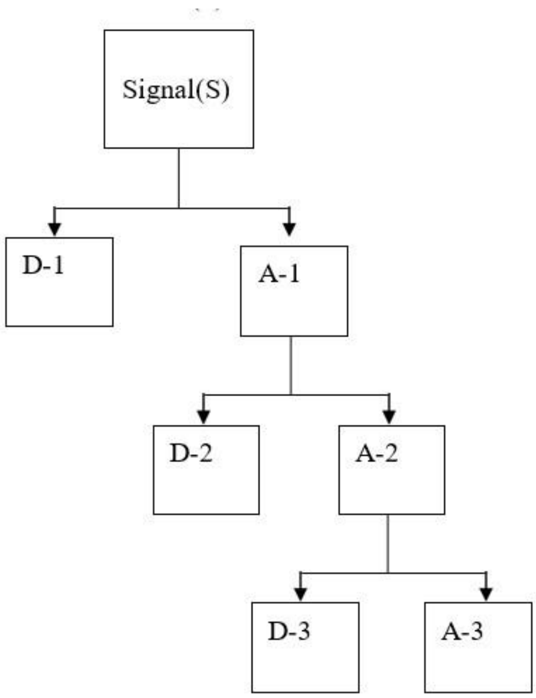


Fig 6: Wavelet Decomposition of signal

The Original Signal S can be reconstructed with the help of A3, D3, D2 and D1

So, it is clear that with the help of equation (4) and equation (1), (2), (3) one can decompose any original signal sequences in to wavelet decomposition. And with this wavelet decomposition again the wavelet is reconstructed.

The number of samples in next decomposition level is half as compared to previous stage. Suppose The original signal S has N samples then A1 and D1 will have N/2 Samples and A2 and D2 will have N/4 Samples. So, wavelet transform is highly suitable for the analysis of local behavior of the signal such as spikes or discontinuities. Because at the point of discontinuity the frequencies changes very fast only for a very little time so by choosing suitable time scale we can also study or analysis these sudden changes.

So, 4 level wavelet decomposition is applied to the two records. As a result, 5 coefficients (a, d1, d2, d3, d4) for each record i.e., healthy and unhealthy. The dimensions of each coefficient record obtained by performing 4 level wavelet decomposition is 4098×100

3.4 FEATURE EXTRACTION

In proposed model mean, variance, skewness, kurtosis, max_svd & entropy_svd are extracted. They are briefly explained as below.

1. **Mean:** Mean is defined as the ratio of the sum of all the observations in the data to the total

number of observations. This is also known as Average. Thus, mean is a number around which the entire data set is spread.

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \ldots \ldots \ldots 3.1$$

where n = total number of observations

$\sum x_i$ = sum of the observations

2. **Variance**: Variance measures how far are data points spread out from the mean. A high variance indicates that data points are spread widely and a small variance indicates that the data points are closer to the mean of the data set. It is calculated as

$$\text{Varience} = \frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})^2 \ldots \ldots \ldots 3.2$$

where N = Total number of observations

$\bar{x}$ = mean of the observations

3. **Skewness:** Skewness measure of asymmetry in a probability distribution is defined by Skewness. It can either be positive, negative or undefined.

- Positive Skew — This is the case when the tail on the right side of the curve is bigger than that on the left side. For these distributions, mean is greater than the mode.
- Negative Skew — This is the case when the tail on the left side of the curve is bigger than that on the right side. For these distributions, mean is smaller than the mode.

The most commonly used method of calculating Skewness is

$$\mathbf{skewness} = \frac{3(\text{Mean} - \text{Median})}{\text{std Deviation}} \ldots \ldots \ldots 3.3$$

If the skewness is zero, the distribution is symmetrical. If it is negative, the distribution is Negatively Skewed and if it is positive, it is Positively Skewed. The below graph shown in Figure 6″ the 3 types of skewness

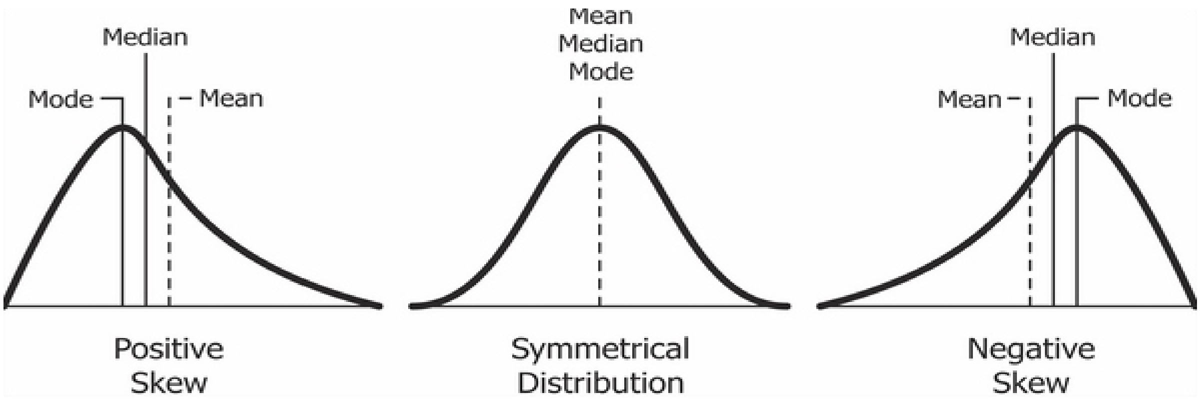


FIG 6′′: Different types of skewness

4. **Kurtosis**: Kurtosis describes the whether the data is light tailed (lack of outliers) or heavy tailed (outliers present) when compared to a normal distribution. There are three kinds of Kurtosis:

$$kurt = \frac{\mu_4}{\sigma^4} \ldots\ldots\ldots 3.4$$

$$where\ \mu_4 = fourth\ cenral\ moment$$

$$\sigma^4 = standard\ deviation$$

- Mesokurtic — This is the case when the kurtosis is zero, similar to the normal distributions.
- Leptokurtic — This is when the tail of the distribution is heavy (outlier present) and kurtosis is higher than that of the normal distribution.
- Platykurtic — This is when the tail of the distribution is light (no outlier) and kurtosis is lesser than that of the normal distribution.

5. **Max_Svd:** The Singular Value Decomposition (SVD) of a matrix is a factorization of that matrix into three matrices. It has some interesting algebraic properties and conveys

important geometrical and theoretical insights about linear transformations. It also has some important applications in data science. Here we consider maximum value of the svd values.

Mathematics behind SVD

The SVD of mxn matrix A is given by the formula:

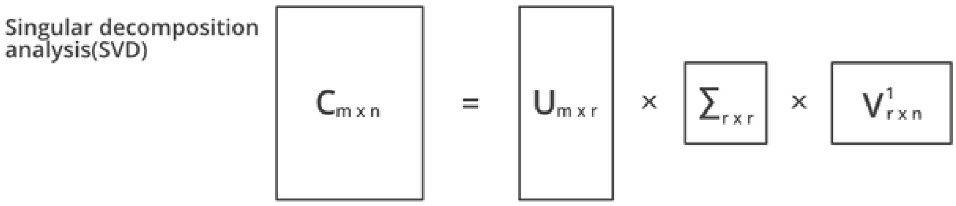
$$A = UWV^{T} \dots\dots\dots 3.5$$

where:

U: *mxn* matrix of the orthonormal eigenvectors of .

V^T: transpose of a *nxn* matrix containing the orthonormal eigenvectors of A^{T}A.

W a *nxn* diagonal matrix of the singular values which are the square roots of the eigenvalues of $A.A^T$



6. **Entropy_Svd:** Shannon entropy (or just *entropy*) is a measure of uncertainty (or variability) associated with random variables. It was originally developed to weigh the evenness and richness of animal and plant species (Shannon, 1948). Exactly how you calculate entropy is very field specific. For example, you wouldn't calculate nutrition in the same way you calculate entropy in thermodynamics. However, all formulas are based on Shannon's original metric, which was calculated as follows:

$$H = \sum_{i=1}^{N} p_i \ln\left(\frac{1}{p_i}\right) \ldots\ldots\ldots 3.6$$

Where:

- H = Shannon Entropy,
- P_i = fraction of population composed of a single species i,
- ln = natural log,
- S = how many species encountered,
- Σ = summation of species 1 to S

3.5 CLASSIFICATION OF ABNORMAL EEG

The Machine Learning Framework consists of 7 classification models

They are:

1. Support Vector Machine (SVM)
2. K-Nearest Neighbor (KNN)
3. Decision Tree
4. Random Forest
5. AdaBoost
6. Bagging
7. Logistic Regression

3.5.1 Support Vector Machine (SVM)

Support Vector Machine (SVM) is a supervised machine learning algorithm used for both classification and regression. Though we say regression problems is well its best suited for classification. The objective of SVM algorithm is to find a hyperplane in an N-dimensional space that distinctly classifies the data points. The dimension of the hyperplane depends upon the number of features. If the number of input features is two, then the hyperplane is just a line. If the number of input features is three, then the hyperplane becomes a 2-D plane. It becomes difficult to imagine when the number of features exceeds three. In this case, the new variable y is created as a function of distance from the origin. A non-linear function that creates a new variable is referred to as kernel. Below figure 7 shows the principle of SVM classifier.

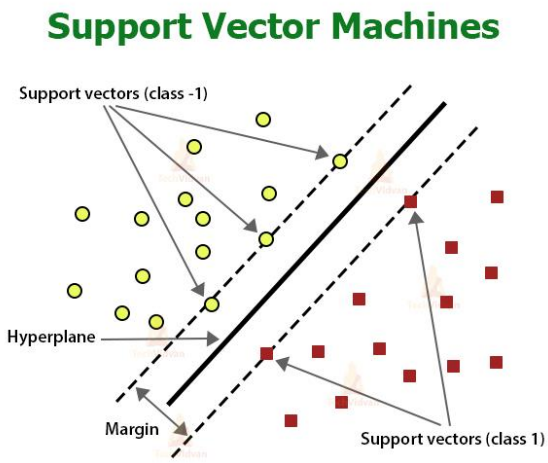


Fig 7: SVM Classifier [12]

SVM Kernel:

The SVM kernel is a function that takes low dimensional input space and transforms it into higher-dimensional space, i.e. it converts not separable problem to separable problem. It is mostly useful in non-linear separation problems. Simply put the kernel, it does some extremely complex data transformations then finds out the process to separate the data based on the labels or outputs defined.

3.5.2 K-Nearest Neighbor (KNN)

K-Nearest Neighbors is one of the most basic yet essential classification algorithms in Machine Learning. It belongs to the supervised learning domain and finds intense application in pattern recognition, data mining and intrusion detection. It is widely disposable in real-life scenarios since it is non-parametric, meaning, it does not make any underlying assumptions about the distribution of data (as opposed to other algorithms such as **Gaussian mixture model (GMM)**, which assume a Gaussian distribution of the given data). We are given some prior data (also called training data), which classifies coordinates into groups identified by an attribute. Below Figure 8 shows the principle of KNN classifier.

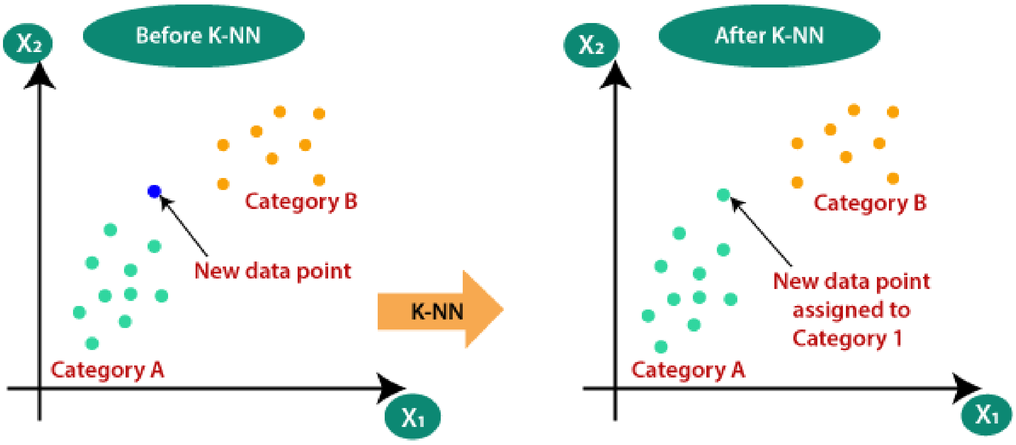


Fig 8: KNN Classifier [13]

3.5.3 Decision Tree

Decision tree is the most powerful and popular tool for classification and prediction. A Decision tree is a flowchart like tree structure, where each internal node denotes a test on an attribute, each branch represents an outcome of the test, and each leaf node (terminal node) holds a class label.

A tree can be *"learned"* by splitting the source set into subsets based on an attribute value test. This process is repeated on each derived subset in a recursive manner called *recursive partitioning*. The recursion is completed when the subset at a node all has the same value of the target variable, or when splitting no longer adds value to the predictions. The construction of decision tree classifier does not require any domain knowledge or parameter setting, and therefore is appropriate for exploratory knowledge discovery. Decision trees can handle high dimensional data. In general decision tree classifier has good accuracy. Decision tree induction is a typical inductive approach to learn knowledge on classification. Below figure 9 shows the general structure and principle of Decision Tree.

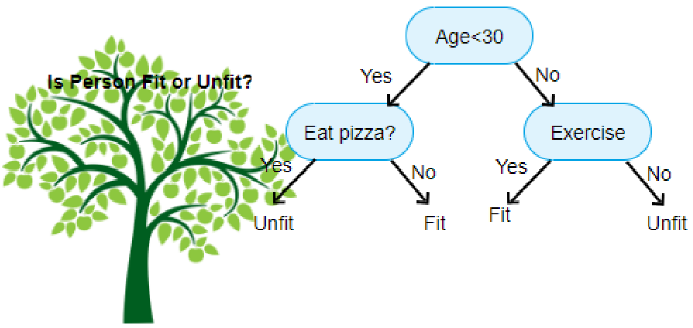


Fig 9: Decision tree for the concept of fitness [16]

3.5.4 Random Forest

A Random Forest is an ensemble technique capable of performing both regression and classification tasks with the use of multiple decision trees and a technique called Bootstrap and Aggregation, commonly known as **bagging**. The basic idea behind this is to combine multiple decision trees in determining the final output rather than relying on individual decision trees.

Random Forest has multiple decision trees as base learning models. We randomly perform row sampling and feature sampling from the dataset forming sample datasets for every model. This part is called Bootstrap. The principle of Random Forest is as shown in below.

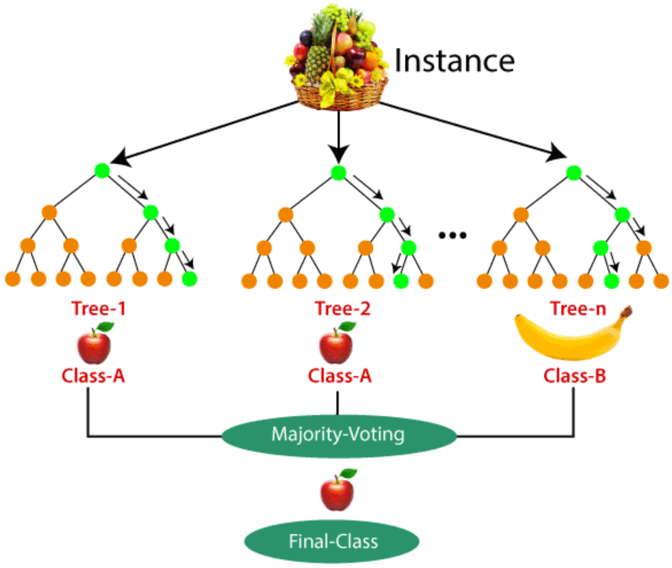


Fig 10: Random Forest Algorithm [14]

3.5.5 AdaBoost

AdaBoost was the first really successful boosting algorithm developed for the purpose of binary classification. AdaBoost is short for Adaptive Boosting and is a very popular boosting technique that combines multiple “weak classifiers” into a single “strong classifier”. The most common algorithm used with AdaBoost is decision trees with one level that means with Decision trees with only 1 split. These trees are also called **Decision Stump.** The general structure of stump is as shown below figure 11. Here black color represents the main node and green color represent the leaf node

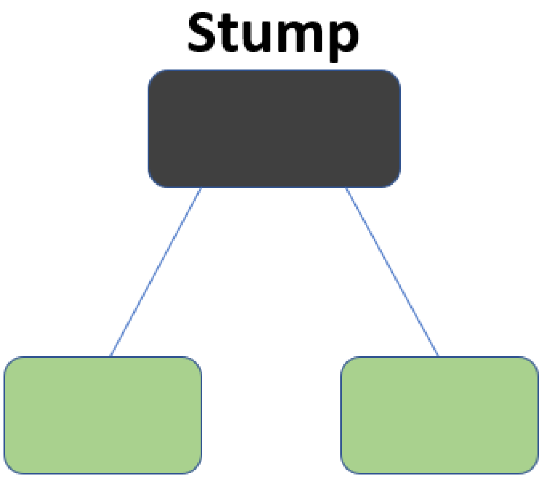


Fig 11: Stump [15]

What this algorithm does is that it builds a model and gives equal weights to all the data points. It then assigns higher weights to points that are wrongly classified. Now all the points which have higher weights are given more importance in the next model. It will keep training models until and unless a lower error is received. The below figure 12 shows the ensemble technique which is the principle of AdaBoost

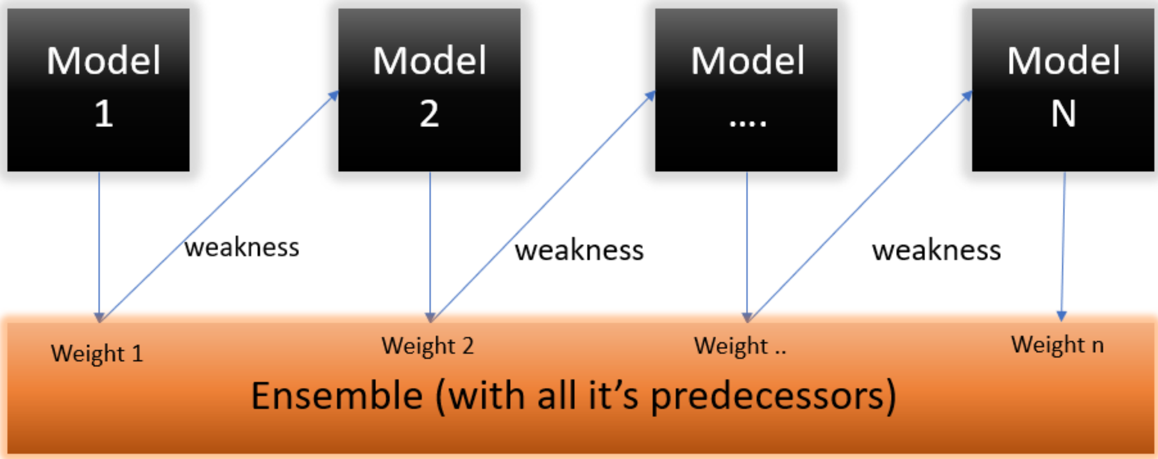


Fig 12: AdaBoost Algorithm [15]

3.5.6 Bagging

A Bagging classifier is an ensemble meta-estimator that fits base classifiers each on random subsets of the original dataset and then aggregate their individual predictions (either by voting or by averaging) to form a final prediction. Such a meta-estimator can typically be used as a way to reduce the variance of a black-box estimator (e.g., a decision tree), by introducing randomization into its construction procedure and then making an ensemble out of it.

Each base classifier is trained in parallel with a training set which is generated by randomly drawing, with replacement, N examples (or data) from the original training dataset – *where N is the size of the original training set*. Training set for each of the base classifiers is independent of each other. Many of the original data may be repeated in the resulting training set while others may be left out. Bagging reduces overfitting (variance) by averaging or voting, however, this leads to an increase in bias, which is compensated by the reduction in variance though. The principle of Bagging technique is as shown below figure 13

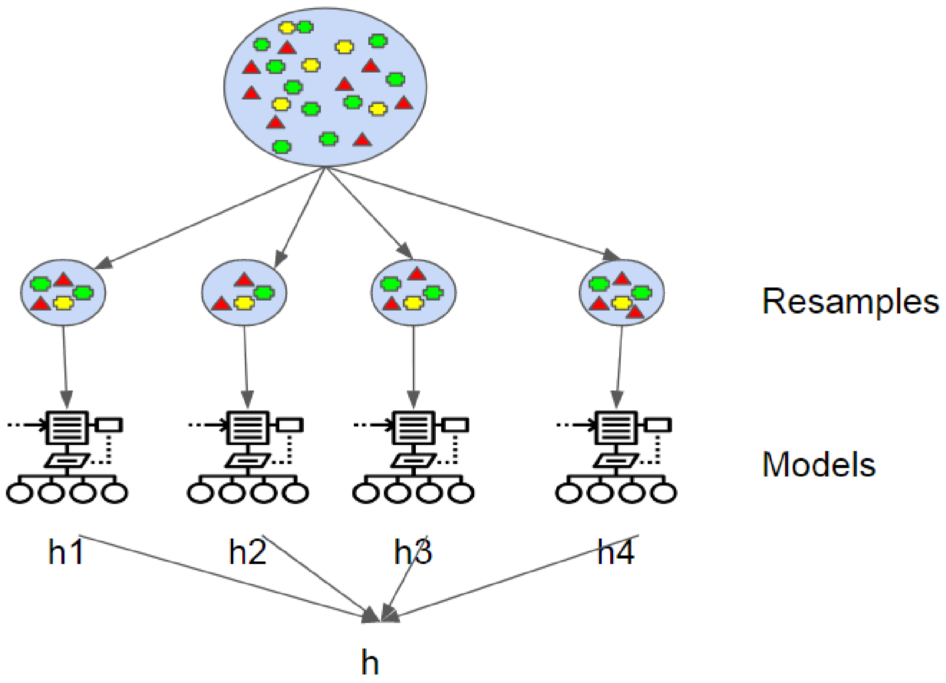


Fig 13: Bagging Algorithm [17]

Logistic Regression

Logistic regression is basically a supervised classification algorithm. In a classification problem, the target variable (or output), y, can take only discrete values for a given set of features Contrary to popular belief, logistic regression IS a regression model. The model builds a regression model to predict the probability that a given data entry belongs to the category numbered as “1”. Just like Linear regression assumes that the data follows a linear function, Logistic regression models the data using the sigmoid function shown below.

$$g(z) = \frac{1}{1 + e^{-z}} \ldots\ldots\ldots 3.6$$

Logistic regression becomes a classification technique only when a decision threshold is brought into the picture. The setting of the threshold value is a very important aspect of Logistic regression and is dependent on the classification problem itself. The principle of Logistic

Regression is as shown below figure 14

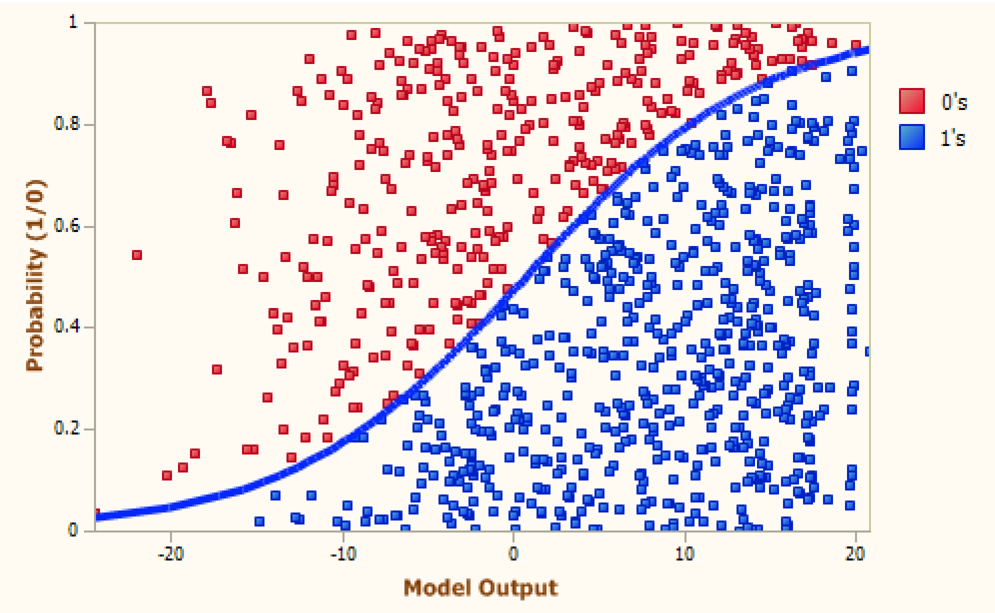


Fig 14: Logistic Regression Classifier [18]

Now feature data of each coefficient records of both healthy and unhealthy obtained after feature extraction are split into two sets (train & test) in a ratio such that 80% as train data and 20% as test data. By using the train data, each classifier is built and predict the abnormality in EEG signal. Performance of the model is calculated by comparing the predicted results with test record results

CHAPTER 4

RESULTS AND DISCUSSION

RESULTS AND DISCUSSION

Different classification models are built based on train data and predict the results and compare them with test data using some parameters they are

1. ACCURACY
2. PRECISION
3. RECALL
4. F1 SCORE
5. CONFUSION MATRIX
6. ROC CURVE

First of all, whenever we perform classification predictions there are four types of outcomes that could occur:

1. **True Positive (TP)**
2. **True Negative (TN)**
3. **False Positive (FP)**
4. **False Negative (FN)**

True Positive:

A **true positive** is an outcome where the model correctly predicts the positive class.

True Negative:

A **true negative** is an outcome where the model correctly predicts the negative class.

False Positive:

A **false positive** is an outcome where the model incorrectly predicts the positive class.

False Negative:

A **false negative** is an outcome where the model incorrectly predicts the negative class.

Confusion matrix

It is a performance measurement for machine learning classification problem where output can be two or more classes. It is a table with 4 different combinations of predicted and actual values. It is extremely useful for measuring Recall, Precision, Specificity, Accuracy, and most importantly AUC-ROC curves.

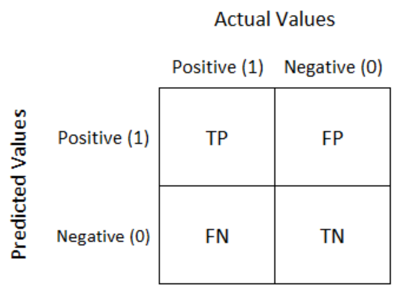


Fig 15: Confusion Matrix [19]

ACCURACY:

Accuracy is the most intuitive performance measure and it is simply a ratio of correctly predicted observation to the total observations. One may think that, if we have high accuracy then our model is best. Yes, accuracy is a great measure but only when you have symmetric datasets where values of false positive and false negatives are almost same. Therefore, you have to look at other parameters to evaluate the performance of your model.

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \dots \dots \dots 4.1$$

PRECISION:

Precision is the ratio of correctly predicted positive observations to the total predicted positive observations. High precision relates to the low false positive rate.

$$Precision = \frac{True\ positive}{True\ Positive + False\ Positive} \dots \dots \dots 4.2$$

RECALL:

The ratio of correct positive predictions to the total predicted positives.

$$Recall\ = \frac{True\ positive}{True\ Positive + False\ Negatives} \dots \dots \dots 4.3$$

F1 SCORE:

F1 Score is the weighted average of Precision and Recall. Therefore, this score takes both false positives and false negatives into account. Intuitively it is not as easy to understand as accuracy, but F1 is usually more useful than accuracy, especially if you have an uneven class distribution.

$$F1 = 2 * \frac{Precision * recall}{precision + recall} \dots \dots \dots 4.4$$

ROC CURVE An **ROC curve** (**receiver operating characteristic curve**) is a graph showing the performance of a classification model at all classification thresholds. This curve plots two parameters:

- True Positive Rate
- False Positive Rate

True Positive Rate (**TPR**) is a synonym for recall and is therefore defined as follows:

$$TPR = \frac{TP}{TP + FN} \dots \dots \dots 4.5$$

False Positive Rate (**FPR**) is defined as follows:

$$FPR = \frac{FP}{TN+FP} \dots \dots \dots 4.6$$

The classification report & confusion matrix of various classifiers (SVM, KNN, Decision Tree, Random Forest, AdaBoost, Logistic Regression & Bagging) in the ML Framework are shown below.

1. SVM

The performance of the svm classifier built using the train data in terms of confusion matrix and ROC curve as shown in the below figures 16 & 17

```
print("SVM Classifier")
report(clf)
```

```
SVM Classifier
              precision    recall  f1-score   support

           0       0.90      0.49      0.64       184
           1       0.06      0.38      0.10        16

    accuracy                           0.48       200
   macro avg       0.48      0.43      0.37       200
weighted avg       0.83      0.48      0.60       200

/usr/local/lib/python3.7/dist-packages/sklearn/utils/deprecation.py:87:
  warnings.warn(msg, category=FutureWarning)
/usr/local/lib/python3.7/dist-packages/sklearn/utils/deprecation.py:87:
  warnings.warn(msg, category=FutureWarning)
```

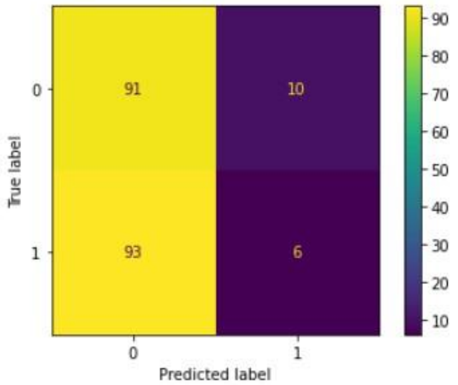


Fig 16: Classification Report and Confusion Matrix

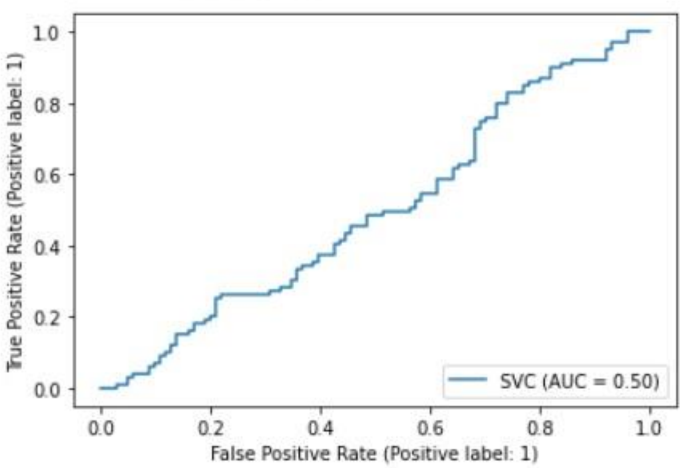


Fig 17: ROC Curve

2. Decision Tree

The performance of the Decision Tree classifier built using the train data in terms of confusion matrix and ROC curve as shown in the below figures 18 & 19

```
Decision Tree Classifier
              precision    recall  f1-score   support

           0       0.59      0.69      0.64        89
           1       0.71      0.62      0.66       111

    accuracy                           0.65       200
   macro avg       0.65      0.65      0.65       200
weighted avg       0.66      0.65      0.65       200

/usr/local/lib/python3.7/dist-packages/sklearn/utils/deprecation
  warnings.warn(msg, category=FutureWarning)
/usr/local/lib/python3.7/dist-packages/sklearn/utils/deprecation
  warnings.warn(msg, category=FutureWarning)
```

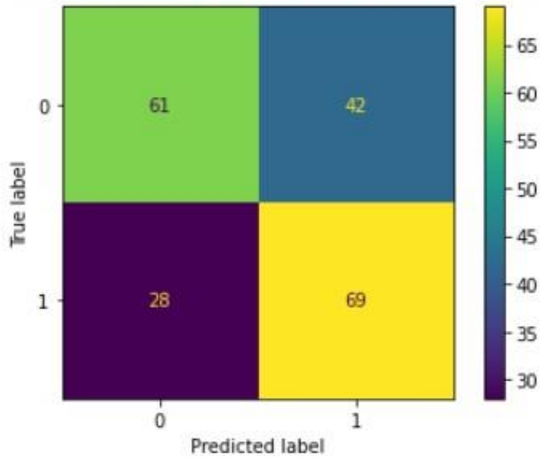


Fig 18: Classification Report and Confusion Matrix

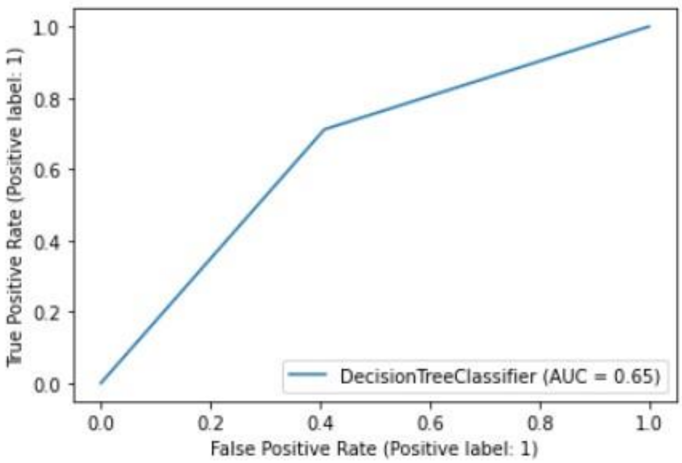


Fig 19: ROC Curve

3. Random Forest

The performance of the Random Forest classifier built using the train data in terms of confusion matrix and ROC curve as shown in the below figures 20 & 21

```
print("Random Forest Classifier")
report(ran)
```

```
Random Forest Classifier
              precision    recall  f1-score   support

           0       0.61      0.72      0.66        88
           1       0.74      0.64      0.69       112

    accuracy                           0.68       200
   macro avg       0.68      0.68      0.67       200
weighted avg       0.68      0.68      0.68       200

/usr/local/lib/python3.7/dist-packages/sklearn/utils/deprecati
  warnings.warn(msg, category=FutureWarning)
/usr/local/lib/python3.7/dist-packages/sklearn/utils/deprecati
  warnings.warn(msg, category=FutureWarning)
```

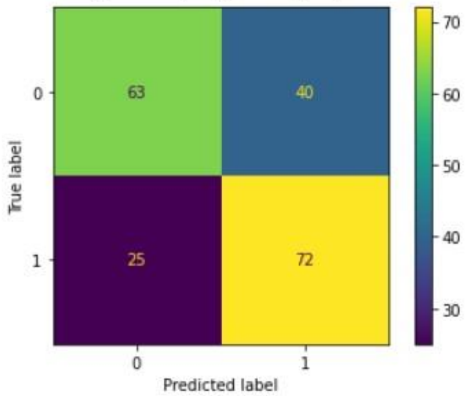


Fig 20: Classification Report and Confusion Matrix

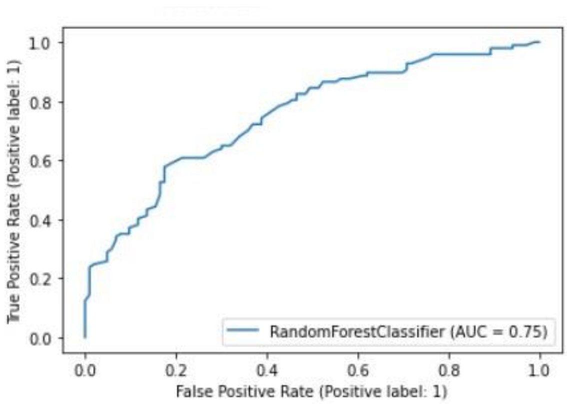


Fig 21: ROC Curve

4. AdaBoost

The performance of the AdaBoost classifier built using the train data in terms of confusion matrix and ROC curve as shown in the below figures 22 & 23

```
Adaboost
              precision    recall  f1-score   support

           0       0.54      0.61      0.57        92
           1       0.63      0.56      0.60       108

    accuracy                           0.58       200
   macro avg       0.59      0.59      0.58       200
weighted avg       0.59      0.58      0.59       200

/usr/local/lib/python3.7/dist-packages/sklearn/utils/deprecation.
  warnings.warn(msg, category=FutureWarning)
/usr/local/lib/python3.7/dist-packages/sklearn/utils/deprecation.
  warnings.warn(msg, category=FutureWarning)
```

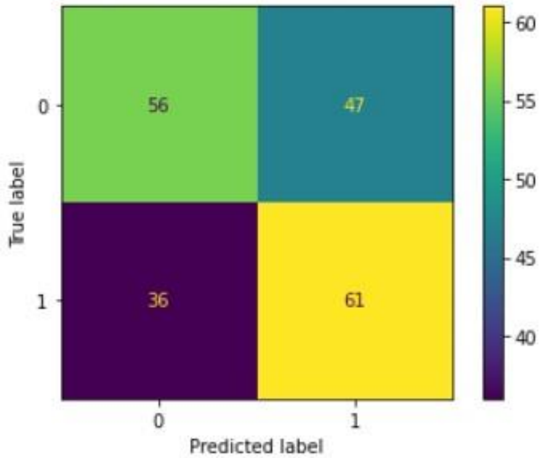


Fig 22: Classification Report and Confusion Matrix

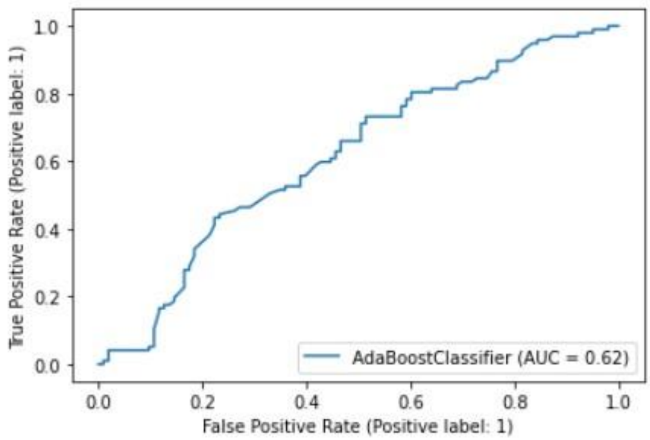


Fig 23: ROC Curve

5. KNN

The performance of the KNN classifier built using the train data in terms of confusion matrix and ROC curve as shown in the below figures 24 & 25

```
KNN
              precision    recall  f1-score   support

           0       0.46      0.59      0.51        80
           1       0.66      0.53      0.59       120

    accuracy                           0.56       200
   macro avg       0.56      0.56      0.55       200
weighted avg       0.58      0.56      0.56       200

/usr/local/lib/python3.7/dist-packages/sklearn/utils/deprecation.
  warnings.warn(msg, category=FutureWarning)
/usr/local/lib/python3.7/dist-packages/sklearn/utils/deprecation.
  warnings.warn(msg, category=FutureWarning)
```

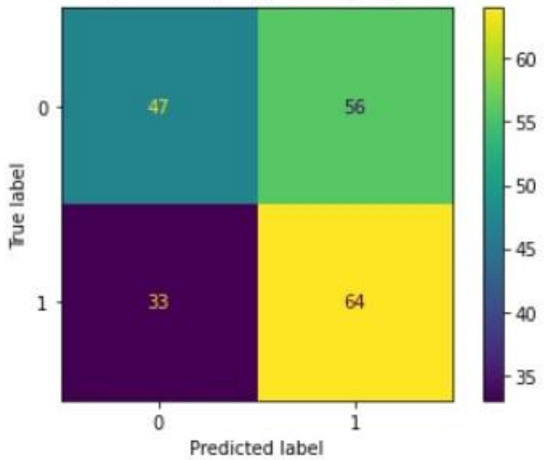


Fig 24: Classification Report and Confusion Matrix

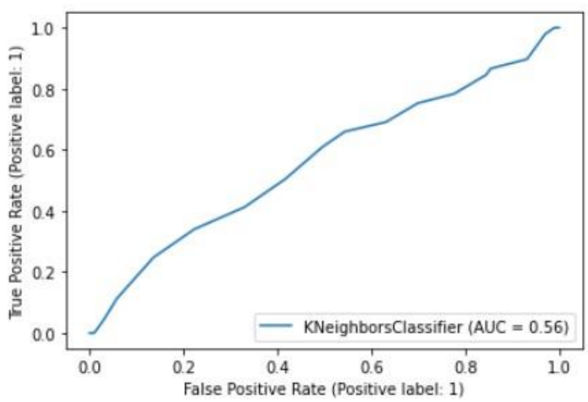


Fig 25: ROC Curve

6. Logistic Regression

The performance of the Logistic Regression classifier built using the train data in terms of confusion matrix and ROC curve as shown in the below figures 26 & 27

```
Logistic Regression
              precision    recall  f1-score   support

           0       0.65      0.49      0.56       136
           1       0.29      0.44      0.35        64

    accuracy                           0.48       200
   macro avg       0.47      0.47      0.45       200
weighted avg       0.53      0.47      0.49       200

/usr/local/lib/python3.7/dist-packages/sklearn/utils/deprecation
  warnings.warn(msg, category=FutureWarning)
/usr/local/lib/python3.7/dist-packages/sklearn/utils/deprecation
  warnings.warn(msg, category=FutureWarning)
```

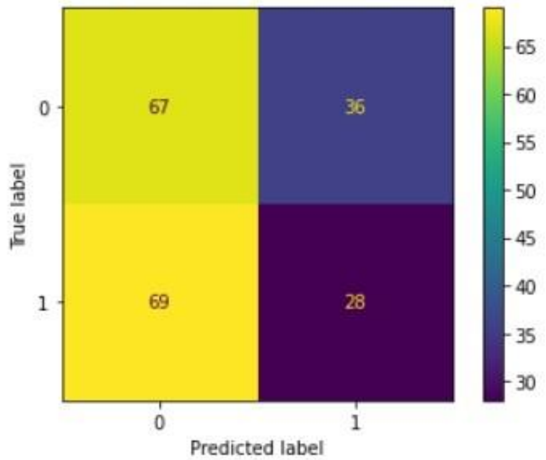


Fig 26: Classification Report and Confusion Matrix

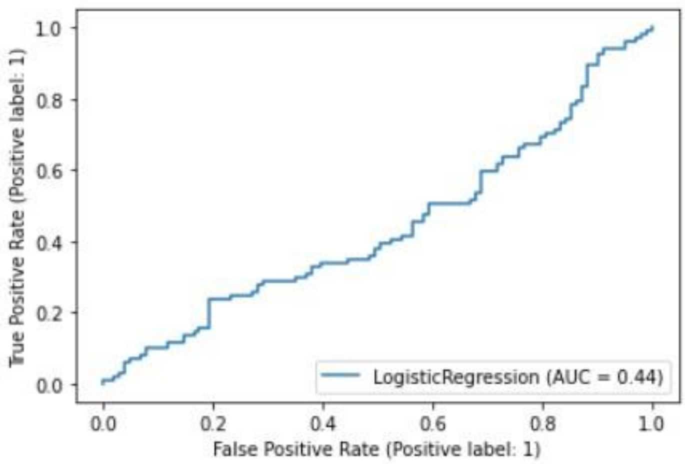


Fig 27: ROC Curve

7. Bagging

The performance of the Logistic Regression classifier built using the train data in terms of confusion matrix and ROC curve as shown in the below figures 28 & 29

```
bagging
              precision    recall  f1-score   support

           0       0.65      0.74      0.69        91
           1       0.75      0.67      0.71       109

    accuracy                           0.70       200
   macro avg       0.70      0.70      0.70       200
weighted avg       0.71      0.70      0.70       200

/usr/local/lib/python3.7/dist-packages/sklearn/utils/deprecation.
  warnings.warn(msg, category=FutureWarning)
/usr/local/lib/python3.7/dist-packages/sklearn/utils/deprecation.
  warnings.warn(msg, category=FutureWarning)
```

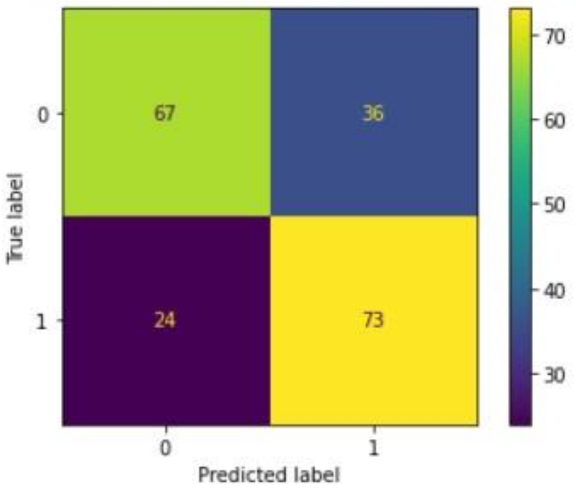


Fig 28: Classification Report and Confusion Matrix

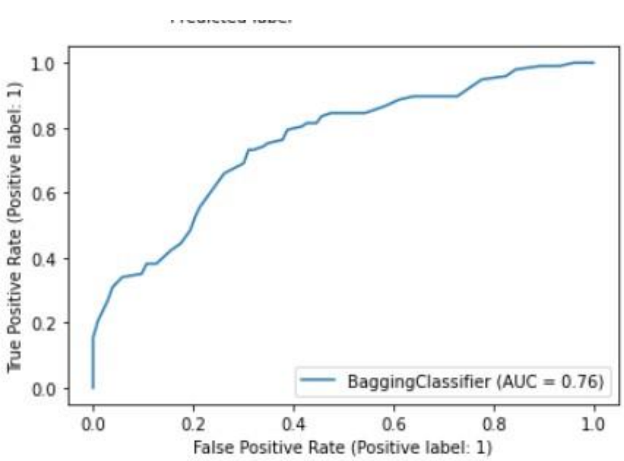


Fig 29: ROC Curve

Table 4.1: Various classifiers performance

CLASSIFIER	ACCURACY (%)	ROC (%)
SVM	48	50
Decision Tree	65	65
Random Forest	68	75
AdaBoost	58	62
KNN	56	56
Logistic Regression	48	44
Bagging	70	76

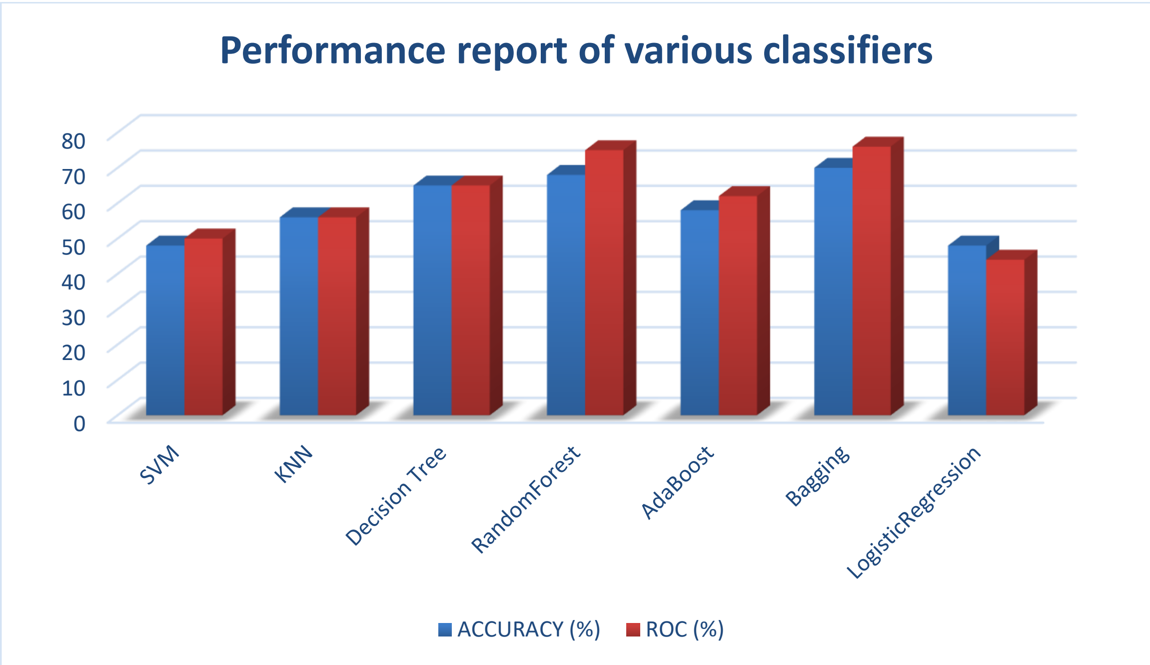


Fig 30: Performance report of the various classifiers

By observing the above table 4.1 and graph shown in figure 30 above Among all classifier bagging classifiers perform well for the abnormal (seizure activity in EEG) classification of EEG as the accuracy of this classifier is 70% and area under ROC curve is 76%

CONCLUSIONS

CONCLUSIONS

A brain is an organ that serves as the center of the nervous system in all vertebrate and most invertebrate animals and it is located in the head, usually close to the sensory organs for senses such as vision. Brain is having more complex neural network as a result brain signals are quite complex to study. Visual inspection of the EEG signal by observing a change in frequency or amplitude in long-duration signals is an arduous task for the clinicians. It may lead to an erroneous classification of EEG abnormalities. So, a model is proposed such that the main objective of model is to aid the radiologist in diagnosis of abnormalities of the EEG. The proposed model consists a ML Frame work with various classification algorithms to classify normal and abnormal EEG which represent seizure activity. This ML frame work is fed with EEG data from the Bonn dataset then this signal undergoes 4 level wavelet decomposition from the decomposed data features like mean, variance, skewness, kurtosis, max_svd & entropy_svd are extracted and fed to various classification algorithms in the ML framework model i.e., SVM, KNN, random forest, AdaBoost, logistic regression, Boosting, Decision tree. Out of all these Classification algorithms in ML framework bagging classifier perform well with 70% accuracy and 76% AUC.The propose model can reduce the workload on the radiologist who work on detecting seizures visually and manually on long-duration EEG signal.ML framework model requires less computation than CNN model. Since our framework model is simple and resourceful it can used in Realtime for the detection of abnormality (seizure) in EEG signal. Thus, ML framework will rectify the misclassification of abnormalities in EEG signals.

This machine learning frame work can be used to other abnormalities like sleep disorder, neural diseases like Parkinson's, Alzheimer's, autism, etc. Research work can be done to build a wearable device which predict the spontaneous abnormalities in the brain like seizures etc.

References

[1] Sharma, Manish, Sohamkumar Patel, and U. Rajendra Acharya. "Automated detection of abnormal EEG signals using localized wavelet filter banks." *Pattern Recognition Letters* 133 (2020): 188-194.

[2] Singh, Gurwinder, Manpreet Kaur, and Dalwinder Singh. "Detection of epileptic seizure using wavelet transformation and spike based features." *2015 2nd International Conference on Recent Advances in Engineering & Computational Sciences (RAECS)*. IEEE, 2015.

[3] Liu, Yinxia, et al. "Automatic seizure detection using wavelet transform and SVM in long-term intracranial EEG." *IEEE transactions on neural systems and rehabilitation engineering* 20.6 (2012): 749-755.

[4] Phadikar, Souvik, Nidul Sinha, and Rajdeep Ghosh. "Automatic Eyeblink Artifact Removal From EEG Signal Using Wavelet Transform With Heuristically Optimized Threshold." *IEEE Journal of Biomedical and Health Informatics* 25.2 (2020): 475-484.

[5] Peachap, Atemangoh Bruno, and Daniel Tchiotsop. "Epileptic seizures detection based on some new Laguerre polynomial wavelets, artificial neural networks and support vector machines." *Informatics in Medicine Unlocked* 16 (2019): 100209.

[6] Faust, O., Acharya, U.R., Adeli, H. and Adeli, A., 2015. Wavelet-based EEG processing for computer-aided seizure detection and epilepsy diagnosis. *Seizure*, *26*, pp.56-64.

[7] Sifuzzaman, M., M. Rafiq Islam, and M. Z. Ali. "Application of wavelet transform and its advantages compared to Fourier transform." (2009).

[8] https://www.chihealth.com

[9] Cabrera, Dalton H. Bermudez. "Implementation of Machine Learning Algorithm to Exploit Information from Multimodal fMRI/EEG Fused Image Data."

[10] http://www.vakhospital.com

[11] https://towardsdatascience.com/the-wavelet-transform-e9cfa85d7b34

[12] https://techvidvan.com/tutorials/svm-in-r/

[13] https://www.javatpoint.com/k-nearest-neighbor-algorithm-for-machine-learning

[14] https://www.section.io/engineering-education/introduction-to-random-forest-in-machine-learning/

[15] https://www.mygreatlearning.com/blog/adaboost-algorithm/

[16] https://www.aitimejournal.com/@akshay.chavan/a-comprehensive-guide-to-decision-tree-learning

[17] http://www.differencebetween.net/technology/difference-between-bagging-and-random-forest/

[18] https://patvin.co.in/jvaa.aspx?iid=107069162-logistic+regression+titanic+python&cid=28

[19] https://towardsdatascience.com/understanding-confusion-matrix-a9ad42dcfd62

[20] M. Sharma, R.B. Pachori, U.R. Acharya, A new approach to characterize epileptic seizures using analytic time-frequency flexible wavelet transform and fractal dimension, Pattern Recognit. Lett. 94 (2017) 172–179

[21] https://www.who.int/news-room/fact-sheets/detail/epilepsy

Appendix

Coding part

https://github.com/19PA1A0465/AI-LAB/blob/master/model_with_approx.ipynb