Machine learning approach for early prediction of Dementia

Mr. Vishal Patil[1]
Assistant Professor, Computer Engineering Department, K. J. Somaiya Institute of Engineering and Information Technology, Mumbai, India
patil.vr@somaiya.edu[1]

Ms. Nisha Vanjari[2],
Assistant Professor, Computer Engineering Department, K. J. Somaiya Institute of Engineering and Information Technology, Mumbai, India
nvanjari@somaiya.edu[2]

Mr. Yogesh Jadhav[3], **Ms. Deepa Parshar**[4]
Assistant Professor, Computer Engineering Department, Amity University, Mumbai, India
yhjadhav@mum.amity.edu[3], <dparasar@mum.amity.edu> [4]

Abstract*—Dementia is a disorder caused by the gradual abnormality and death of brain cells thus progressively rendering the person to be dependent on others for their activities of daily living. With symptoms like muffled speech, memory loss, visual hallucinations, change in concentration and judgement, the paper proposes to forecast dementia in its early stage which has no established cure till date. The aim is to slow down its progress through early and accurate prediction using EEG (Electroencephalography) signals. EEG signals are used to extract meaningful markers to characterise the effects of Dementia on the brain. The paper proposes to extract features from the frequency bands which will be given as an input to Machine Learning algorithm. An application for doctors, EEG technicians and patients will be the user interface to carry out this entire process. This application will serve as a clinical tool which aims at serving as an aid for the doctor's fraternity and not a substitute. It will be in accordance with the clinical code of conduct.

***Keywords*—Dementia, EEG, Machine Learning.**

I. Introduction

Dementia is a state in which daily activities of living get affected. In India, more than 4 million people are suffering from Dementia[1]. Globally, at least 44 million people have some form of Dementia, thus this disease with global health crisis must be addressed.[1] In its early stage, the symptoms are minimal but as the disease progresses the symptoms worsen. Also, since Dementia is incurable, the diagnosis of this disease at its early stage is crucial.

Detection of Dementia at an early stage is essential to slow down its progress. Following core mental functions must be significantly impaired to be considered Dementia: loss of memory, difficulty in interpretation, change in concentration level, muffled speech, behavioural changes and visual hallucinations.

Electroencephalography (EEG) is a clinical symptomatic instrument that straightforwardly reflects cortical neuronal working and hence can be applied to the analysis of dementia [7]. Various parameters exist which can be used for classifying the EEG signals. The following parameters can be used for analysis with respect to the diagnosis: Frequency, Voltage. Dementia affects brain signals associated with the neurons, thus affecting major frequency bands such as beta(>13Hz), alpha (8-12 Hz), theta (4-7 Hz) and Delta (0-4 Hz) from the normal level [7].

According to The World Alzheimer Report 2015 led by King's College London, India is home to 4.1 million cases of Dementia. As per survey, several hospitals lack the facility of dementia detection and thus the disease is detected in a conventional way using MRI, mental status tests (Eye movement, speech, reflexes are considered). Also, due to lack of awareness regarding Dementia in India, there is a need for a diagnostic tool to predict and hence try to prevent it.

EEG brain signals provide the necessary features that are directly related to Dementia diagnosis. Thus, a machine learning approach will help in forecasting the dementia at an early stage by using EEG signals. We have trained the datasets of the Electroencephalogram obtained from healthy and abnormal subjects to extract meaningful markers from the EEG signal of dementia patients which reflect pathological changes.Apart from using EEG signals for diagnosis, we have also made use of other determining factors such muffled speech, visual hallucinations. We have used multiple Machine Learning algorithms such as K Nearest Neighbour, Bayes Theorem and compared their accuracies.

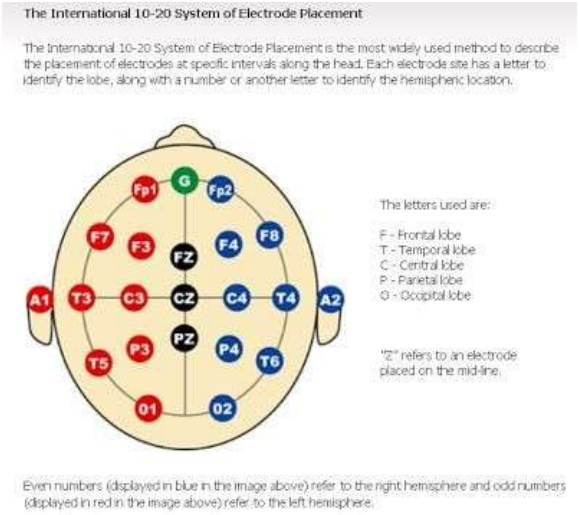

Fig 1.1: EEG Electrode placement [9]

II. EXISTING SYSTEM

Consolidating Single-Channel EEG Measurement and Verbal Fluency Test-A Groundwork for Ambulatory Diagnosis of Dementia [7]

A Verbal Fluency Test (VFT) is a psychological test where the subject matter has to say as many words as possible from a given type in 60 seconds. Accuracy is measured based on the number of right answers. Right now, distinction saw in the EEG examples of a basic subject playing out a straightforward VFT task and an ordinary subject performing hard VFT undertaking will demonstrate comparable attributes to the distinction saw in the EEG designs between a typical subject playing out a simple VFT and potential Alzheimer's patient performing same VFT. This method uses the EEG signals captured during the VFT and checks if there is a possibility of Dementia. The test explains that there is a decrease in the alpha and beta frequency ranges prominently and an increase in the delta and theta frequencies.

EEG Markers for Early detection and characterization of Vascular Dementia during Working Memory tasks[8]

In this paper, Spectral entropy, sample entropy as well as Hurst Exponent is used to reveal markers from the EEG signals on subjects which include 5 patients suffering from vascular dementia (VaD), 15 suffering with stroke-related mild cognitive impairment (MCI) as well as 15 control healthy subjects performing a working memory (WM) task[8]. Independent principal component analysis is used for removing the artifacts from the EEG signals. Spectral Entropy is used for testing the hypothesis that EEG signal in both VaD and MCI slows down in comparison with control subjects. In differentiate, SampEn (Sample Entropy) and H (Hurst Exponent) highlights are utilized for testing the speculation of decrease of the abnormality and multifaceted nature which can be found in VaD and MCI when contrasted with the control subjects[8].

It is passed on that example entropy could be utilized to distinguish VaD (Vascular Dementia) and Hurst Exponent for stroke related Mild Cognitive Impairment recognizable proof. EEG information is gotten utilizing NicoletOne (V32) framework. Straight and non-direct highlights should be acquired from the EEG signals. The paper utilizes Spectral Entropy method to gauge easing back of the EEG signals where as non-direct systems are utilized to decide complex unique information about the cerebrum.

The paper passed on that the EEG exercises showed that the VaD and MCI patients had marginally decreased qualities at all scalp areas and the thing that matters isn't important ($p>0.05$) when Spectral Entropy (SpecEnVaD<SpecEnMCI<SpecEnControl) is used[8]. The paper presumes that Sample Entropy can be utilized for VaD recognition and Hurst Exponent for stroke related MCI distinguishing proof.

A Process of Analysis of EEG Wave Trains in Early Stages of Parkinson's Disease[5]

The technique exhibited utilizes wave trains which are designs in a foundation EEG and investigations these waves in wide recurrence groups like alpha, beta and theta EEG[5]. The examination of the EEG wave trains includes discovery neighborhood maxima, assurance and factual investigation of different properties of the maxima. EEG accounts have been taken from a lot of 18 patients determined to have the primary stage Parkinson's infection where they were told to keep up a particular posture which incites tremor.

The examination uncovers that itemized time-recurrence elements ought to be considered as for alpha and beta groups in order to stay away from any incorrect understandings because of the tails of the alpha waves as an electrical action in the beta groups. Consequently this examination has given a technique to exploring the mind electrical exercises of the patients by investigating the EEG wave trains. The after effects of this investigation are contrasted and the outcomes got from standard examination. Results from standard strategy uncover an important increment of the force unearthly thickness in Parkinson's illness in C3 and C4 terminals in alpha band while a critical reduction in the measure of the wave prepares in similar cathodes in the close by beta groups is uncovered.

Prediction of Brain Diseases using EEG and Speech Signal [6]

The most overwhelming brain diseases are Parkinson's disorder (PD) and Alzheimer's ailments. People developed 50 or above generally experience the evil impacts of Parkinson's sickness [6] and this strategy proposes usage of EEG and talk signals. This technique utilizes discrete wavelet change highlight extraction

Technique. Support Vector Machine algorithm is used for categorization, giving 91.6% accuracy Pre-processing of EEG signals is done using zero low pass filters and 13 features are extracted.

Mel frequency Cepstral Coefficients(MFCC) is very efficient which is used for feature extraction from the speech signals. Minimum distance classifier is used for classifying the speech signals which is used for recognising the voice characteristics of the person. EEG dataset consist of 20 people where 8 suffered from PD, 8 from Alzheimer and 4 were healthy and speech signal dataset consists of 70 people where 30 were detected with Parkinson's, 30 with Alzheimer and 10 people were detected healthy[6].

III. Proposed System Elements

Following Machine Learning algorithms have been used in existing systems:

1. Support Vector Machine (91.6% accuracy)
2. Back Propagation using Neural Networks (85% accuracy)
3. Novelty Detection Algorithm (90% accuracy)
4. K means clustering

Based on our extensive research through IEEE papers and information provided on net, we have used K-means clustering, KNN and Bayes Theorem for classification.

Parameters of EEG

The main frequencies of human EEG are:

1. Alpha: Alpha waves are mainly in the range 7.5-13Hz. The majority of the alpha waves can be observed in adults. The waves are usually perceived during relaxation and when a person's eyes are closed. The waves tend to recede during actions like calculation, thinking and when a person's eyes are open.

2. Beta: The range of Beta waves spans greater than 14 Hz. It is highly observed in patients who are alert or anxious. The waves are absent or reduced in areas of cortical damage.

3. Theta: Theta waves range from 3.5 to 7.5Hz. It appears normal when a person sleeps and is also observed during metabolic encephalopathy. Theta waves have two types- the first is widespread distribution across scalp which indicates decreased alertness, cognitive impairment, dementia and the second is distributed within the frontal midline and it indicates mental alertness and focusing.

4. Delta: The range of such waves is lesser than 3Hz. Delta waves are recorded in the frontal position in adults and in the posterior area in children.

These are the 4 major frequencies observed in the recorded EEG signals for predicting the chance of a person suffering from Dementia.

IV. System Modules

The system consists of 2 basic elements where algorithms are coded in python and Django which has a model-view-controller architecture is used for providing User Interface.

a) Doctor Login and Registration on the portal:
On the backend we have a database where the details of the registered doctors will be stored and they will have the access to the patients records.

b) Analyze Results:
This module contains following sub-modules:

i) Predict if there is any slowing seen in the EEG signals of the patient. This takes input as alpha, beta, theta, delta as the parameters.

ii) Predict the chances of Dementia

- The output of the submodule is the input to this sub-module. The parameters chosen here are: Age, EEG slowing, Memory loss, Difficulty in Interpretation, Visual Hallucination, Behavioural changes and Muffled Speech.

- View Accuracy: It predicts the accuracy of the algorithms used. If the algorithms give high accuracy then the doctor can decide to consider this data for training or he can neglect it.

- Plot Graph: Graphs are plotted in which the current input gets plotted against the training set to provide visual appearance.

c) View Patient's History
This module provides information about a patient stored in the database.

d) Provide Diagnosis
According to the analyzed result, doctors can provide their own diagnosis for the patient.

V. Proposed System

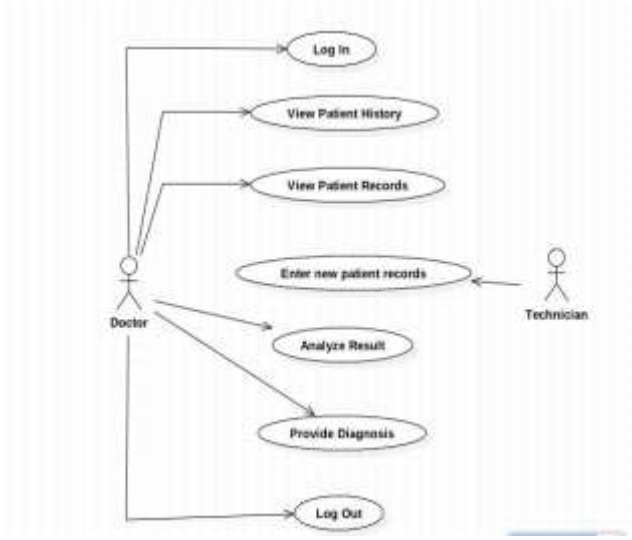

Fig 1.2: Proposed Model Use case diagram[2

The proposed model allows the doctors to login in the system with valid credentials and add, update, delete and analyse the patient records. Our initial dataset consists of EEG signals of 15 patients. The data is preprocessed which leads the way for feature extraction. Here features including frequency bands like alpha, beta, theta and delta as well as wave symmetry are chosen for analysis. The features are classified using following Machine learning algorithms- KNN, Bayes classifier and data analysis show whether there is slowing of EEG signals or not. Then the output of this algorithm is used as a feature along with the major factors which are used by doctors while diagnosing the Dementia patients which include age of the subject, presence or absence of the following observations- memory loss, visual hallucinations, muffled speech, difficulty in interpretations, behavioural changes which are analysed using KNN algorithm and it determines the chances of the subject having Dementia on the basis of the above parameters. Since this system will only aid the doctor in diagnosis, depending on the accuracy of the result for a specific input record, the doctor can decide whether the record should be added to the training data of the system so as to enhance the algorithm used in the system which will help in increasing the accuracy. Also, data is clustered into two groups- patients with slowing of EEG and without of EEG signals and is shown on graph. Whenever a new patient record entered it is also mapped to either of the clusters and displayed with a different marker.

VI. Results

Based on the analyzed graphs, there is decrease in alpha and beta frequency compared to the normal range, which confirms EEG signal slowing, as per our literature survey. Slowing of signals mainly in the left and right temporal region was recorded. By the system implemented by us, we first determine if there is any slowing observed in the EEG records of the patients. Further, the output of this algorithm serves as an input to the next where other parameters like symptoms are added. All the algorithms provide good accuracy of 80% and above.

Screenshots

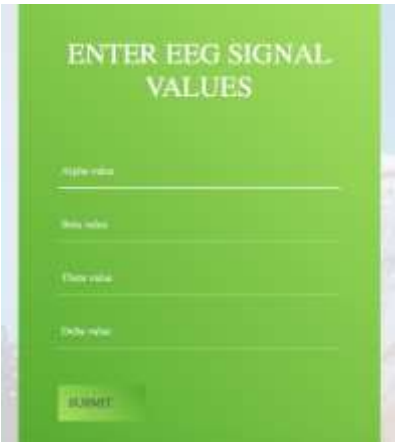

Fig 1.3: Input screen for doctors to enter EEG signal values[3]

Fig 1.4: Input screen for doctors to enter values about commonly observed symptoms[2]

Fig 1.5: Results [5]

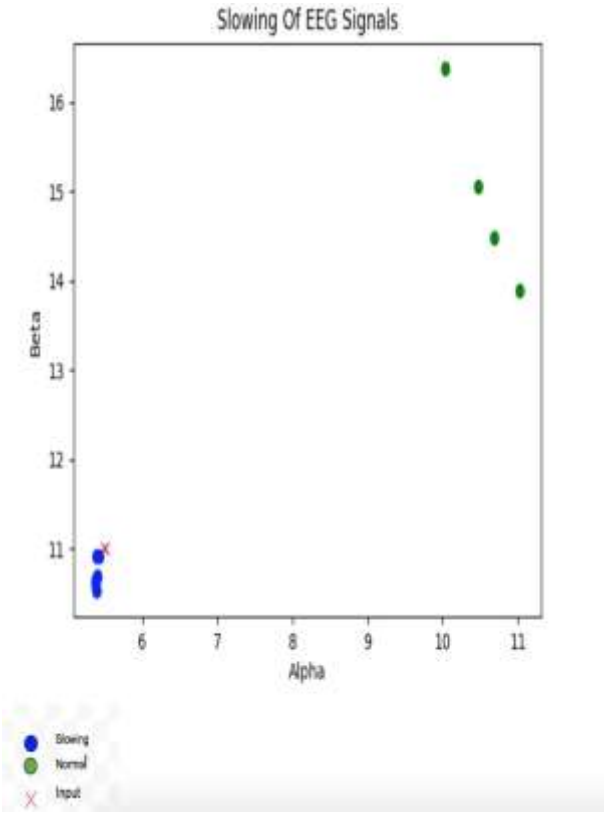

Fig 1.6: Graph displaying two clusters depicting slowing and non-slowing EEG patients.[4]

VII. LIMITATIONS

The proposed model does not guarantee presence of Dementia. It will aid the doctor by telling about the chances and this will help doctor by suggesting whether further analysis is required to confirm its presence.

VIII. CONCLUSION

Spikes indicate an abnormal record, but not necessarily Dementia. Same goes for asymmetric graphs of the left and right hemisphere. Alpha waves appear when relaxing, closing of eyes and disappears while concentrating. Beta waves are dominant in patients who are alert or anxious. Decrease in such features indicate slowing of waves.

The EEG is used as an analyzing tool of diseases, classification and rehabilitation. We can successfully classify between normal healthy people and people suffering from Dementia, by extracting various features from the EEG signals. To implement classification and training, Machine Learning algorithm is used. Hence we can conclude that EEG signals provide the basis for detecting brain diseases and the technique developed shows promising results.

IX. FUTURE SCOPE

The process can be extended for other brain related disorders and further research can be carried out to pinpoint types of Dementia like Vascular, Lewy body, Alzheimer's and Parkinson's disease. Data gathered can used along with big data analysis for predicting people in which sector have more chances of getting diagnosed with Dementia. The project can be coupled with Life Insurance companies that cover these diseases as an annual test. It can be further extended to predict the chances of Dementia for people from different professional sectors, age, gender..

X. REFERENCES

[1] https://www.alz.org/in/dementia-alzheimers-en.asp

[2] M.Teplan "Fundamentals of EEG measurement",Measurement Science Review, Volume 2, Section 2, 2002

[3] Faridah Abd Rehman, Mohd Fauzi Othman, Nurul Aimi Shaharuddin 'Analysis Methods of EEG signals: A review in EEG Application for Brain Disorder' Center for Artificial Intelligence and Robotics, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia 15 October 2014

[4] K.Gomathi, D.Leela, S.Prasad 'Smart System to Recognize EEG Signal for Finding Brain Diseases Using K-Means Clustering' International Journal of Advanced Computer Research December-2013.

[5] Olga S. Sushkova, Alexei A. Morozov, Alexandra V. Gabova 'A Method of Analysis of EEG Wave Trains in Early Stages of Parkinson's Disease', IEEE 2016

[6] Akshata S. Agarwal and Kishori S. Degaonkar 'Detection of Brain Diseases using EEG and Speech Signal' *International Journal of Computer Applications (0975 – 8887) Volume 149 – No.9*, September 2016

[7] Je-Eon Lee, Jaeho Park, Sujeong You 'Combining Single-Channel EEG Measurement and Verbal Fluency Test- A Groundwork for Ambulatory Diagnosis of Dementia' International conference on Ubiquitous Robots and Ambient Intelligence June 28- July 1, 2017

[8] Noor Kamal Al-Qazzaz, Sawal Ali, Shabiul Islam, Siti Anom Ahmad, Javier Escudero 'EEG Markers for Early detection and characterization of Vascular Dementia during Working Memory tasks, IEEE 2016

[9] https://people.ece.cornell.edu/land/courses/ece4760/FinalProjects/s2012/cwm55/cwm55_mj294/