Detection of Different Brain Diseases from EEG Signals Using Hidden Markov Model

Department of Electronics & Communication Engineering, Khulna university of Engineering & Technology, Khulna, Bangladesh
Email: dip.hasin@gmail.com; robi@ece.kuet.ac.bd

Received: 01 June 2019; Accepted: 27 June 2019; Published: 08 October 2019

Abstract—The brain imaging device, Electroencephalography (EEG) provides several advantages over other brain signals like Functional Near-infrared Spectroscopy (fNIRS) and Functional Magnetic Resonance Imaging (fMRI). It is non-invasive and easily applicable. EEG provides high temporal resolution with a low setup cost. EEG signals of several subjects which record electric potential caused by neurons firing in the brain are undergone a Hidden Markov Model (HMM) classification technique. We are particularly interested to detect the brain diseases from EEG signals by an HMM probabilistic model. This HMM model is built with a given initial probability matrix of five different states, namely, epilepsy, seizure, dementia, stroke and normality. The transition probability matrix is updated after each iteration of parameter estimation using Baum-Welch algorithm (B-W algorithm).

Index Terms—Electroencephalography (EEG), Hidden Markov Model (HMM), Baum-Welch algorithm (B-W algorithm), Initial probability matrix, Transition probability matrix.

I. INTRODUCTION

The applications of Electroencephalography (EEG), as a brain signal, is expanding worldwide, mainly for diagnostic purposes and preventive measures of neurological diseases such as epilepsy, stroke, Parkinson’s diseases and others. The continuous monitoring of EEG signals is necessary to observe the electrical activity of the brain to evaluate drug intoxication, trauma and blood flow during surgical procedures. A key advantage of EEG is that it provides higher temporal resolution than other available brain signals as it responds to the cognitive activities of a subject very rapidly (0.5-130 milliseconds).

Hidden Markov Model (HMM) is a non-linear probabilistic classifier and it provides the probability of a given set of series in time domain. As a non-linear classification technique, it shows higher classification accuracy than those of linear classifiers, for example, linear discriminant analysis (LDA) and support vector machine (SVM). In recent years, HMM has been applied

in the various fields of bioinformatics, data mining, pattern recognition, data analysis, wireless networks etc. Some notable works in recent times are protein secondary structure prediction based on a HMM model for data mining [1], offline recognition cursive of Arabic handwritten text without explicit segmentation [2], muscle-computer interface based on HMM state transitions which uses ultrasound sensing [3], action recognition by Gaussian-Mixture HMM (GMM-HMM) model which yields a greater recognition accuracy [4].

HMM has made its mark in different medical, biology and rehabilitation fields; for example, identifying movement states of Parkinsonian patients [5], personal identification system [6], functional brain networks [7], quantification of wheezing for respiratory sound classification [8], single-molecule data analysis in time series which accommodates complications such as drift [9]. Septic shock of critical care patient causes multiple organ failure and eventual death. A HMM model which predicts septic shock for ICU patients [10]. Rehabilitation of a deaf person is done by a DNN-HMM hybrid system for lip-reading and audio visual speech recognition (AVSR) [11], fall detection and real-life home monitoring for senior citizens [12], emotion classification by a combined SVM-HMM classifier to recognize human emotion states based on EEG signals [13]. Already a stacked HMM model has found its applications in robotics for motion intention recognition based on motion trajectories [14].

Other recent and notable studies of HMM include removal of EEG artifact caused by eye blinks [15], seisimocardiography learning based on expectation-maximization algorithm and Viterbi algorithm [16], automatic volcano-seismic events detection based on HMM with state and event duration models [17], neural prostheses to restore efficient communication to people with motor neurological injury [18], obstructive sleep apnea (OSA) detection based on ECG signals [19], 3D catheter tip tracking inside the patient’s 3D vessel tree using 2D X-ray image sequences and a peri-operative 3D rotational angiography (3DRA) [20], palm rehabilitation with supervised learning[21], automatic significant beats extraction in Holter register and feature extraction of ECG signals [22]. Many studies, mostly of functional near-infrared spectroscopy (fNIRS) have successfully
demonstrated the use of LDA, SVM and ANN for BCI [23]. Application of machine learning methods has been conducted to combine the EEG-INIRS signals [24].

We have proposed a methodology in this paper which will classify the EEG signals using HMM technique. The key advantages of HMM are high classification accuracy, provision to add more states etc. as we have found in our research. However, as a brain imaging technique EEG provides lower spatial resolution as compared with INIRS.

II. METHODOLOGY

A. EEG Signal Acquisition System

Electrodes are likely to be placed on the frontal region for the EEG data acquisition. The positions for electrode placements are Frontal polar 1 (Fp1) and Frontal polar 2 (Fp2). For the single channel data acquisition process positive electrode is placed on the Fp2 position, reference electrode is placed on the Fp1 position and negative electrode is placed on A1 (earlobe) position according to international 10-20 electrode system (Fig.1).

Fig.1 Electrode placement system in 10-20 electrode system

B. EEG Signal Conditioning System

Signal conditioning unit has been designed to amplify the EEG signals for further processing (Fig.2). EEG signals picked up using electrodes have a magnitude of around 5-500μV and frequency range of 1-100 Hz, as the most information of EEG signals lie in this range. The integrated single supply instrumentation amplifier is designed based on a typical customized three op-amp approach. It is a high precision, low-power general purpose amplifier. An active high pass filter and an active low pass filter have been designed to operate in the cut-off frequency range of 1-100 Hz.

For removing the 50 Hz power-line interference, a notch filter has been designed for better performance. A post amplifier for further amplification of the EEG signal and a voltage adder circuit has been designed for keeping the signal in positive domain.

C. Hidden Markov Model (HMM)

A hidden Markov model (HMM) is a state machine with two layers, namely state layer and observation layer (Fig.3). An HMM model is defined by three probability matrices- initial, transition and emission probabilities. The initial probability is defined as the probability of each state in the beginning of the Markovian process. The transition probability defines the probability of shifting from one state to another (or the same state) in a particular sequence. The emission probabilities define the probability of a certain observation symbol $x_t$ in state $s_t$.

\[
p(q_n|q_{n-1}, q_{n-2}, \ldots, q_2, q_1) = p(q_n|q_{n-1})
\]  

(1)

In standard HMMs, at a certain time, say $t$, all the needed information of the previous states (before $t$) would be known by using a discrete hidden state. For a sequence of $n$ elements, say $Q$, where, $Q = \{q_1, q_2, q_3, \ldots, q_n\}$ and

\[
p(q_n|q_{n-1}, q_{n-2}, \ldots, q_2, q_1) = p(q_n|q_{n-1})
\]  

(1)

Equation (1) is called the first order Markov assumption, where, the probability of observation of a symbol, $q_n$ at time $n$ depends only on observation of $q_{n-1}$ at time $(n-1)$.

An HMM model is specified by

a) the (hidden) state sequence $S = \{s_1, s_2, s_3, \ldots, s_n\}$,

b) the observation sequence $X = \{x_1, x_2, x_3, \ldots, x_n\}$,

c) a set of parameters, $\theta = (\pi, A, B)$ where, $\pi$ = a vector of initial probabilities of each state, $A = $ a matrix of transition probabilities, $B = $ a vector of probabilistic output matrix

Mathematically,

\[
\begin{align*}
\pi_i &= p(s_t = t = 0) \\
\alpha_{ij} &= p(s_t = t + 1| s_t = t) \\
b_i(x) &= p(O_t = x| s_t = t)
\end{align*}
\]

(2) \quad (3) \quad (4)

Where,

$x = $ observed signal value;

$s = $ state of the HMM model

The output function $B$ is a Gaussian density function.
\[ n(O, \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \]  

(5)

Where,

- \( O \) = sequence of observation;
- \( \mu \) = mean of the sequence;
- \( \sigma \) = standard deviation of the sequence.

However, the input state sequences are unknown in our current study.

For a given \( \theta \) we have:

\[ p(x_1, x_2, x_3, ... x_n | \theta) = p(x_1 | \theta) ... p(x_n | \theta) \]  

(6)

Since each observation value \( x_i \) is independent.

For each sequence \( x \):

\[ p(x | \theta) = \sum_s p(x, s | \theta) \]  

(7)

The sum is taken over all hidden state paths \( s \).

We are intended to include the Baum-Welch algorithm in the HMM model for the parameter estimation when state paths are unknown (Fig.4). This is basically an Expectation Maximization algorithm in the iterative process of HMM. When the states are known, we can simply count. However, when the states are unknown, the 'counting' process is little trickier. In such cases the 'average' process is employed. For each edge, the 'average' number is computed for all possible pairs of given states [25].

D. Baum-Welch (B-W) Algorithm

Actually the Baum-Welch training is an expectation maximization algorithm suitable for HMM classification when the input states are unknown. We start with some initial values of transition and emission probabilities which define prior values of \( \theta \). Then we use an iterative algorithm which attempts to replace \( \theta \) by \( \theta^* \) such that

\[ p(x | \theta^*) > p(x | \theta) \]  

(8)

Where, \( x \) is the observation state;

- \( \theta = \) Initial parameters \((\pi, A, B)\);
- \( \theta^* = \) Updated parameters \((\pi, A, B)\);

A Markov chain over a set of (hidden) states, and for each state \( s \) and observable symbol \( x \), an emission probability \( p(X_i = x | S_i = s) \). An HMM model is defined by the parameters: \( a_{kl} \) (transition probabilities) and \( e_k(b) \) (emission probabilities) for all states \( k \), \( l \) and all symbols \( b \) as shown in Fig.5. The B-W algorithm will be operated in two steps, namely, the E step and the M steps.

The E Step

For each edge, \( s_{i-1} \rightarrow s_i \), we compute the average number of “\( k \) to \( l \)” transitions, for all possible pairs \((k, l)\) over this edge. Then for each \( k \) and \( l \), we take \( A_{kl} \) to be the sum over all edges. When we have \( n \) independent input sequences \((X_1, ... , X_n)\) then \( A_{kl} \) is given by,

\[ A_{kl} = \frac{1}{n} \sum_{i=1}^{n} \sum_{s_{i-1}=1}^{L} p(s_{i-1} = k, s_i = l, x^i | \theta) \]  

(9)

Where, \( n \) = No. of observation sequence \((X_1, ... X_n)\):

\[ p(x^i) = \text{Probability of observation of a particular symbol, } x^i; \]

\( L \) = No. of total edges, \( s_{i-1} \rightarrow s_i \).

Expected no. of symbol emissions for state \( k \) and each symbol \( b \), for each \( i \), compute the expected no of times that \( X_i = b, E_k(b) \).

\[ E_k(b) = \frac{1}{n} \sum_{i=1}^{n} \sum_{s_i=k} p(x_1, ... x_n, s_i = k) \]  

(10)
The M Step

New values of $a_{kl}$ and $e_k(b)$ after each iteration define $\Theta^*$ and are computed using $A_{kl}$ and $E_k(b)$ from the E step respectively:

$$a_{kl} = \frac{\lambda_{kl}}{\sum \lambda_{kl}} \quad \text{and} \quad e_k(b) = \frac{E_k(b)}{\sum \lambda_{kl}}$$

(11)

E. Data Description

The EEG signals of 130 subjects whose ages are in the range of 10-100 years have been collected for this study from the Temple University Hospital (TUH) EEG Corpus- a big data resource for automated EEG interpretation [26]. The EEG signals used in this study, have been recorded using several generations of NicoletTM EEG recording technology. The number of channels vary between 24 and 36 channels sampled at 250 Hz frequency. The acquired signals are digitized using 16 bit A/D converter [26]. Each EEG file contains about 20 Mbytes of data stored in an European Data Format (EDF+) file format. Later, EDF+ format has been converted to cell data [27]. These EDF files contain a header with 24 unique fields that contains patient’s information and signal’s conditions such as amplitudes, frequency, no of channels etc. Selected fields of this header is shown in Table 1.

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Version Number</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>Patient ID</td>
<td>TUH123456789</td>
</tr>
<tr>
<td>4</td>
<td>Gender</td>
<td>M</td>
</tr>
<tr>
<td>6</td>
<td>Date of Birth</td>
<td>57</td>
</tr>
<tr>
<td>8</td>
<td>Firstname Lastname</td>
<td>TUH123456789</td>
</tr>
<tr>
<td>11</td>
<td>Startdate</td>
<td>01-MAY-2010</td>
</tr>
<tr>
<td>13</td>
<td>Study Number/ Tech. ID</td>
<td>TUH123456789/TAS X</td>
</tr>
<tr>
<td>14</td>
<td>Start Date</td>
<td>01.05.10</td>
</tr>
<tr>
<td>15</td>
<td>Start time</td>
<td>11.39.35</td>
</tr>
<tr>
<td>16</td>
<td>Number of Bytes in Header</td>
<td>6400</td>
</tr>
<tr>
<td>17</td>
<td>Type of Signal</td>
<td>EDF+C</td>
</tr>
<tr>
<td>19</td>
<td>Number of Data Records</td>
<td>207</td>
</tr>
<tr>
<td>20</td>
<td>Dur. of a Data Record (Secs)</td>
<td>1</td>
</tr>
<tr>
<td>21</td>
<td>Number of Signals in a Record</td>
<td>24</td>
</tr>
<tr>
<td>27</td>
<td>Signal Prefiltering</td>
<td>HP: 1.000Hz LP: 70.0Hz N: 60.0</td>
</tr>
<tr>
<td>28</td>
<td>No. of Signal Samples/Rec.</td>
<td>230</td>
</tr>
</tbody>
</table>

6-way classification is focused on the collected EEG signals:

a) **SPSW:**

SPSW stands for Spike and/or Sharp Wave which demonstrates epileptiform transients. Patients with epilepsy have this type of waveform.

b) **PLED:**

The full form of PLED is Periodic Lateralized Epileptiform Discharges which shows EEG abnormalities such as repetitive spike or sharp wave discharges over one hemisphere at almost fixed time intervals.

c) **GPED:**

GPED stands for Generalized Periodic Epileptiform Discharges which provide both short and long interval periodic diffuse discharges along with suppression burst patterns according to interval between discharges. Triphasic waves are included in this class.

d) **ARTF:**

ARTF means artifacts that represent noises in the EEG signals due to equipment, environment or movement of the subjects.

e) **EYEBL:**

EYEBL means eye blinks which is often be confused with a spike.

f) **BCKG:**

BCKG means background which includes all other signals.

The first three classes describe those information which are required in manual interpretation, periodic and occur across the channel. The last three classes are necessary for background modelling which attempts to model the signal temporal evolution.

F. Experiment

In this experiment, 5 states, namely, Epilepsy, Seizure, Stroke, Dementia and Normality are defined with the following transition probability from one state to the next (Fig.6). As we see, the transition probabilities are all equal initially due to the fact that all states are equiprobable. It indicates that a subject may belong to any of the 5 states equally likely depending on his/her EEG signal.

![Fig.6. State transition diagram of the HMM](image-url)
updated after each iteration as given in Equation 11. In proposed methodology, the HMM parameters (transition and emission matrix) are undergone repetitive iterations until a certain threshold is met. This threshold is determined by Equation 8 which is the outcome of expectation maximization algorithm. The initial probability for each of the 5 states is same (0.2). It means subject no= 1 may belong to any of the 5 states as defined above. The states of the next subjects, subject no = 2, 3, 4, 5,…. up to the last subject, depends on the states of their exact previous subject as first order Markov assumption is applied (Equation 1). The states of the subjects are determined from the observation sequence. The algorithm and flowchart used in this study are represented here where the steps are followed sequentially.

Step 1: Define number of states and the HMM parameter θ (π, A, B).

Step 2: Collection of EEG signals from the subjects using 24 to 36 channels

Step 3: Sample the EEG signals at sampling frequency = 250 Hz and represent each sample using 16 bits.

Step 4: Run the HMM model with the initial value of θ (π, A, B) [Equation 2,3,4].

Step 5: Update θ to θ ∗ after each iteration using the B-W algorithm [Equation 9,10.11].

Step 6: Check the threshold \( p(\chi|θ ∗) > p(\chi|θ) \) [Equation 8] after each iteration. If it satisfies then collect hidden and observation sequences. If not satisfies then go to Step 4.

III. Result Analysis

The task to adjust the HMM model parameter, Θ to maximize the likelihood of the observation sequence is known as the training phase [27]. The B-W algorithm is used to address this problem. This study is aimed to classify the EEG signals to detect 4 different brain diseases and the normality condition of 130 subjects. The performance results are shown in Table 2.

Table 2. Performance analysis of the proposed algorithm on 130 subjects

<table>
<thead>
<tr>
<th>Classifier</th>
<th>States</th>
<th>No. of subjects</th>
<th>No. of PD Detected</th>
<th>Sensitivity</th>
<th>Overall accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>HMM</td>
<td>Epilepsy</td>
<td>44</td>
<td>44</td>
<td>100%</td>
<td>92.31%</td>
</tr>
<tr>
<td></td>
<td>Dementia</td>
<td>10</td>
<td>8</td>
<td>80%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Stroke</td>
<td>25</td>
<td>24</td>
<td>96%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Seizure</td>
<td>41</td>
<td>37</td>
<td>90.24%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Normal</td>
<td>10</td>
<td>7</td>
<td>70%</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Performance of other studies on 31 subjects

<table>
<thead>
<tr>
<th>Classifier</th>
<th>States</th>
<th>No. of Subject</th>
<th>No. of PD Detected</th>
<th>Sens. (%)</th>
<th>O.A. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HMM</td>
<td>PD</td>
<td>15</td>
<td>14</td>
<td>93.3</td>
<td>90.32</td>
</tr>
<tr>
<td></td>
<td>Healthy</td>
<td>16</td>
<td>2</td>
<td>86.6</td>
<td>90.32</td>
</tr>
<tr>
<td>LS-SVM</td>
<td>PD</td>
<td>15</td>
<td>13</td>
<td>80%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Healthy</td>
<td>16</td>
<td>1</td>
<td>70%</td>
<td></td>
</tr>
</tbody>
</table>

PD= Parkinson’s Disease O.A.= Overall Accuracy Sens.=Sensitivity

In Table 3, the performance of two other classification studies using HMM model [28] and LS-SVM method [29] has been shown. Two statistical parameters, namely, sensitivity (the percentage of correctly detected subjects on their respective states) and accuracy (the overall percentage of correctly detected subjects) are used for the evaluation of different classification schemes. The proposed method provides better performance than the previous studies on several reasons. There are 5 different states and 130 subjects are examined in this study (Table 2) compared to 2 states and 31 subjects on other studies (Table 3). As the HMM is actually a probabilistic distribution method, the accuracy and sensitivity increase with an increase in number of subjects. Moreover, the proposed method in Table 2 is suitable for large number of subjects rather than the two schemes of Table 3. The methodology of Table 2 is suitable for large scale diagnosis purposes. The sensitivity of epilepsy (100%) and stroke (96%) are higher than both the studies of Table 3 (93.33% and 86.67%). The proposed algorithm discusses the most occurring 4 brain diseases whereas both the studies of Table 3 discuss only 1 disease, Parkinson’s disease. So, proposed method is likely to be more suitable for detection of different brain diseases in a large scale. However, the lower sensitivity of dementia and normality is due to lower number of subjects. The overall accuracy is 92.31% in Table 2 which is larger than both the methods (90.32%) showed in Table 3.
IV. CONCLUSION

The classification of biomedical signals is required for clinical diagnosis of the patients, different machine learning techniques, brain-computer interfaces (BCI) etc. We aimed to classify EEG signals to separate the patients with 4 specified neurodegenerative diseases from those of healthy ones. Three main steps constitute the current study, namely, obtaining EEG data, definition of states & HMM parameters and classification using HMM. The iteration process continues until the condition of maximum likelihood of estimation is met. The investigation of performances shows that the proposed method (Table 2) is much more efficient for classification and interpretation of brain diseases than those of previously studied (Table 3). However, there are certain limitations of EEG data. EEG signal has low spatial resolution (about 10cm) and it is susceptible to motion artifacts. In this study we do not correlate the age and gender of the subjects with the EEG classification. However, previous studies showed that the effect of gender on usual locomotion patterns is not considerable [30].

V. FUTURE WORK

Our future work targets a combined EEG-fNIRS scheme to classify the signals to detect brain diseases. A combination of EEG & fNIRS provides both high temporal & high spatial resolution unlikely when they are applied separately. This work is intended to apply for a large number of subjects in a practical medical field.

REFERENCES

Detection of Different Brain Diseases from EEG Signals Using Hidden Markov Model


Authors’ Profiles

Md. Hasin R. Rabbani was born on 2nd August, 1995 in Naogaon; a northern district of Bangladesh. Currently, he is working as a Teaching Assistant (TA) and pursuing his MS degree at the department of Electronics & Communication Engineering (ECE) in Khulna University of Engineering & Technology (KUET), Khulna. He completed his B.Sc.Engg.(ECE) from KUET in 2018. His major research interests include biomedical signal processing and classification, detection of diseases from biomedical signals in a non-invasive way etc.

Sheikh Md. Rabiul Islam received the B.Sc.in Engg. (ECE) from Khulna University, Khulna, Bangladesh in December 2003, and M.Sc. in Telecommunication Engineering from the University of Trento, Italy, in October 2009 and Ph.D. from University of Canberra, Australia, in 2015. He joined as a Lecturer in the department of Electronics and Communication Engineering of Khulna University of Engineering & Technology, Khulna, in 2004, where he is joined an Assistant Professor in the same department in the effect of 2008. He joined Associate Professor in the same department. Now he joined in this department as Professor since 2018. He has published 3 book chapters, 27 Journal and 28 International conferences. His research interests include biomedical Engineering, Bio-signal & image processing, VLSI for signal processing. He is an IEEE member.