UNIVERSITY OF CALIFORNIA,
IRVINE

Detection of Traumatic Brain Injury Using a Standard Machine Learning Pipeline in Mouse and Human Sleep Electroencephalogram

THESIS

submitted in partial satisfaction of the requirements
for the degree of

MASTER OF SCIENCE

in Electrical Engineering

by

Manoj Vishwanath

Thesis Committee:
Professor Hung Cao, Chair
Distinguished Professor Nikil Dutt
Professor Amir M. Rahmani

2021

DEDICATION

To

my parents, Mr. Y. S. Vishwanath & Mrs. H. N. Suryakala, for their constant support, and belief in my dreams.

my committee members for their constant guidance and encouragement.

Professor A. G. Ramakrishnan, Dr. T. V. Ananthapadmanabha and Assistant Professor Sastry V. Ramachandrula who introduced me to this exciting and evolving field of EEG and BCI and mentored me in crucial aspects of scientific research.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to thank Associate Professor Miranda M. Lim and Dr. Carolyn E. Jones from VA Portland Health Care System, Portland and Oregon Health Science University, Portland for curating the mice and human dataset and lending their constant guidance.

I would like to thank Professor Ramesh Srinivasan from Cognitive Sciences, School of Social Sciences, UCI for sharing his domain knowledge in EEG processing and providing valuable inputs throughout this work.

This work is supported by the NSF CAREER Award #1917105 (H.C.), the NIH R44 #OD024874 (H.C.), the setup fund from the Henry Samueli School of Engineering at UC Irvine (H.C.), VA Biomedical Laboratory Research & Development (BLRD) Career Development Award (CDA) #IK2BX002712 and VA Clinical Science Research Development (CSRD) Merit Review Award I01 CX002022.

ABSTRACT OF THE THESIS

Detection of Traumatic Brain Injury Using a Standard Machine Learning Pipeline in Mouse and Human Sleep Electroencephalogram

By

Manoj Vishwanath

Master Of Science in Electrical Engineering

University of California, Irvine, 2021

Professor Hung Cao, Chair

Traumatic Brain Injury (TBI) is a highly prevalent and serious public health concern. TBI is defined as an alteration in brain functioning or brain pathology initiated by external impacts, such as blunt trauma, penetrating objects, or blast waves which can cause a wide range of functional short- or long-term changes affecting thinking, sensation, language, and emotion, and perhaps most prominently, sleep. Most cases of TBI are mild in nature, yet some individuals may develop following-up persistent disability. The pathophysiologic causes for those with persistent post-concussive symptoms are most likely multifactorial and the underlying mechanism is not well understood. Currently, there are no prognostic markers to predict individuals who are most at risk. Thus, novel approaches to the precise detection and prognostication of mTBI is of utmost importance.

The sleep electroencephalogram (EEG) provides a direct window into neuronal activity during an otherwise highly stereotyped behavioral state and represents a promising quantitative measure for TBI diagnosis and prognosis. With the ever-evolving domain of machine learning, deep convolutional neural networks, and the development of better architectures, these approaches hold promise to solve some of the long entrenched challenges of personalized medicine for uses in recommendation systems and/or in health monitoring systems. In par-

ticular, advanced EEG analysis to identify putative EEG biomarkers of neurological disease could be highly relevant in the prognostication of mild TBI, an otherwise heterogeneous disorder with a wide range of affected phenotypes and disability levels.

In this work, we investigate the use of various machine learning techniques on a cohort of mice and human subjects with sleep EEG recordings from overnight, in-lab, diagnostic polysomnography (PSG) from human subjects and 24 hours recording from mice subjects. An optimal scheme is explored for the classification of TBI versus non-TBI control subjects. The results are promising with an accuracy of 95% in mice and 75% in humans. We are thus confident that, with additional data and further studies, we would be able to build a generalized model to detect TBI accurately, not only via attended, in-lab PSG recordings, but also in practical scenarios such as EEG data obtained from simple wearables in daily life.

Chapter 1

Introduction

Traumatic brain injury (TBI) is defined as an alteration in brain functioning or brain pathology initiated by external impacts, such as blunt trauma, penetrating objects, and/or blast waves. TBI results in physical brain damage, including tearing injuries of white matter, hematomas, or cerebral edema [68]. Consequently, it leads to a cascade of metabolic events which can cause a secondary brain damage possibly due to the generation of free radicals, inflammatory responses, calcium-mediated damage, mitochondrial dysfunction, to name a few. Falls are the cause of nearly half of the TBI related hospitalization [30]. Between 2000 and 2019, more that 400,000 TBI cases have been reported in U.S. service members [27]. Based on the severity of the injury, the effects may last upto few days or entire lifetime. Currently severity of TBI is assessed clinically based on Glasgow Coma Scale (GCS) which is a highly observer dependent and a qualitative measure [110]. The person is assessed based on his/her ability to perform certain actions which is shown in Table.1.1.

TBI can cause a wide range of functional short- or long-term changes affecting thinking, sensation, language, emotion, and perhaps most prominently, sleep [34][103]. About 75% of TBIs that occur each year are concussions or mild TBI (mTBI) [35]. Lack of consensus

Response	Scale	Score/Points
Eye opening response	Open spontaneously	4
	Open to sound	3
	Open to pressure	2
	No eye opening	1
Verbal response	Oriented	5
	Confused conversation	4
	Words discernible	3
	Incomprehensible sounds	2
	No verbal response	1
Motor response	Obeyes command	6
	Movement to stimulus	5
	withdraws from pain	4
	Abnormal flexion	3
	Extensor response	2
	No response	1
Mild TBI = 13-15 points	Moderate TBI = 9-12 points	Severe TBI = 3-8 points

Table 1.1: Glasgow Coma Scale

regarding what constitutes mTBI adds to the complication of the under-diagnosis of the disease. Not all instances of mTBI result in persistent disability, and currently there are no prognostic markers to predict individuals who are most at risk. Investigations show that electroencephalogram (EEG) returns to normal baseline within few minutes or days from the time of injury [58]. Thus, novel approaches to the precise detection and prognostication of mTBI is of utmost importance.

EEG is a non-invasive electrophysiological recording of the neuronal activity of the brain that reflects a summation of synchronous activity of a population of neurons [84]. As a result, this electrophysiological recording is space-averaged by volume conduction and does not have high spatial resolution and EEG channels are often highly correlated spatially. However, it measures modulations of synaptic and action potential fields at high temporal resolution. EEG can be broadly divided into two categories: spontaneous potentials such as sleep rhythms and evoked potentials which are time-locked responses to external stimulus. EEG has been used to study various neurological conditions in clinical applications such

as epilepsy [1], sleep disorders, stroke, Alzheimer's disease [63], and topics of interests in cognitive sciences such as sensor and auditory pathways, memory, motor processes, and general intelligence. Use of quantitative EEG (QEEG) in detection and classification TBI have been discussed in detail in features selection section. Some of the major advantages of EEG over other methods to study functioning of the brain are its low cost, lesser need for highly trained clinicians and bulky equipment, thereby being more suited for real-time monitoring. However it has its challenges. EEG is a non-stationary signal with low signal-to-noise ratio and is highly variable across subjects which makes it difficult build generalize models for EEG analysis. Fig.1.1 shows an example of typical EEG recording set-up with a 256 electrode system. Fig.1.2 shows an epoch of human EEG signal. To overcome the above problems signal processing pipelines with domain specific knowledge is often used. With the advent of advanced machine learning (ML) and deep learning (DL) concepts, better automated approaches with better generalization capabilities have been developed [101].

Figure 1.1: An example of EEG Recording set-up

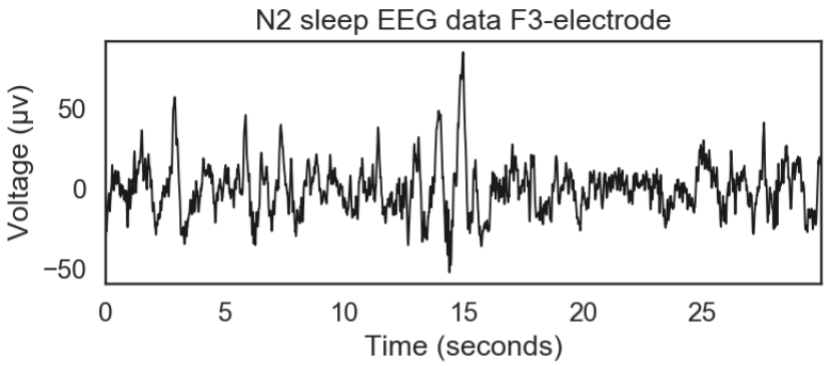

Figure 1.2: EEG signal

Machine learning (ML) and deep learning (DL) are sub-fields of artificial intelligence. Machine learning (ML) [102] refers to the study of computer algorithms that focus on use of statistical methods and data to improve automatically through experience. Rule based ML are characterised by identifying set of rules that represent the knowledge learned by the algorithm. It is dependent on user to determine the set of input features based on domain knowledge. Unlike ML, deep learning automatically extracts relevant feature set from the given data input. ML models can be broadly divided into supervised learning in which the model has access to pre-labeled data, unsupervised learning in which the dataset being used is unlabeled, semisupervised learning in which a small subset of labeled dataset guide classification and feature extraction from a larger, unlabeled dataset, and reinforcement learning that works on the basis of "rewards/punishments" system, offering feedback to the algorithm to learn from its own experiences by trial and error. Detailed description on the different algorithms used in this work are discussed in following sections. In this work we investigate relevant EEG features and various supervised algorithms formulating an efficient pipeline to detect mTBI in mouse and humans sleep EEG.

The following chapters of this thesis have been organised as follows: chapter 2 discusses the data acquisition procedure in mice and human subjects. Chapter 3 and 4 give a detailed description on the pre-processing and feature extraction steps carried out on EEG respectively. Chapter 5 and 6 discusses some of the important and essential steps which have to be taken

before feeding the features on to a machine learning model such as feature normalization and selection techniques. Chapter 7 and 8 sheds light on working of different rule based machine learning algorithms which will be used for classification in this work and their evaluation metrics. Chapter 9 summarizes the pipeline used for mouse and human EEG classification. Finally, chapter 10 and 11 discusses the results obtained and future directions.

Chapter 2

Data Acquisition

2.1 Mouse data

Male C57BL6 mice (Jackson Laboratories, Bar Harbor, ME) 4–6 months of age were used for the experiments. Experimental collection was performed as detailed in [134]. Surgical implantation of microdialysis probes, EEG and electromyography (EMG) wires was performed. EEG and EMG electrodes soldered to a mini-connector for polysomnographic recordings were implanted. A custom-built head stage amplifier (Washington University Electronics Shop, St. Louis, MO), in series with a recording cable distal to the implant was used to eliminate external electrical noise. Following a 2-week recovery on the morning of day 3, left craniotomy and electromagnetic controlled cortical impact (CCI) or left craniotomy only (sham) was performed ipsilateral to the implanted probes. The mice were subjected to CCI with a flat metal tip impounder driven at a velocity of 5 m/sec by an electromagnetic device producing a moderately severe contusion to the cortex [14]. Signals were electronically saved to a file for the offline analysis of sleep states. EEG and EMG activity was assessed using a P511K AC pre-amplifier, digitized with a DigiData 1440A Data Acquisition System,

and recorded using pClamp 10.2. EEG/EMG records were scored automatically using sleep scoring software (SleepSign; Kissei Comtec Co., Ltd., Nagano, Japan) into 4-sec epochs as wake, REM, and non-REM on the basis of standard criteria of rodent sleep [99], and then over-scored manually by visual inspection and corrected when appropriate by a single investigator blinded to intervention (M.M. Lim). On day 5, the animals were sacrificed. Previous results from the same dataset leading up to the work presented in this thesis can be found in [126][125][28].

2.2 Human data

All participants provided informed consent under VA Portland Health Care System (VAPORHCS) Institutional Review Board approval (MIRB #3641). Participants were consented upon referral to the VAPORHCS Sleep Clinic between May 2015 and November 2016 (n = 370). Subjects with in-lab, overnight polysomnography (PSG; n=337) were included in the initial study population. Participants were excluded if they had an apnea-hypopnea index of fifteen or greater (n = 126). Participants who met criteria for PTSD (n=38) were then age-matched to n = 38 non-PTSD controls. Subjects in the data repository underwent in-lab overnight polysomnography (PSG) using Polysmith (NihonKohden, Japan). Six scalp electrodes were placed at F3, F4, C3, C4, O1, and O2 per the 10–20 system of EEG placement. This is shown in Fig.2.1. Following the conclusion of the study, an American Academy of Sleep Medicine (AASM)-accredited polysomnographic technician manually performed standard sleep staging analysis for each 30-second epoch duration according to the standard clinical criteria. Each 30-second epoch of data was scored as one of the five sleep stages (Awake [W], Rapid Eye Movement [REM], non-REM [NREM] stages N1, N2, and N3). Staging of each PSG was additionally validated by a board-certified sleep physician blinded to PTSD status. In the event that there was non-convergence between scorers, the

board-certified sleep physician made the final call. The details of the procedure can be found in [79].

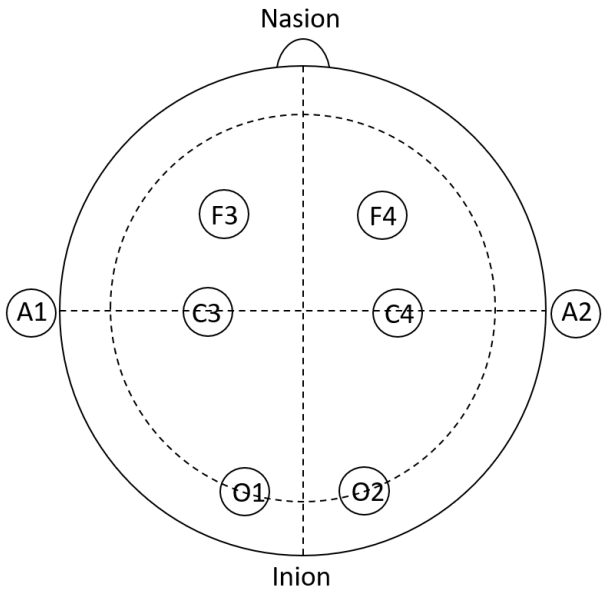

Figure 2.1: Electrode position

Chapter 3

Preprocessing of EEG Signals

Preprocessing is an essential step in the process of EEG analysis. Numerous algorithms and preprocessing pipelines have been developed to tackle the problem of artifact rejection in electrophysiological data. Each of these algorithms has certain strengths and deal with certain aspect of artifact rejection better than the others. To name a few LARG [7] and MARA [135] are ICA-based pipelines whereas ASR_5* and ASR_10* are based on artifact subspace reconstruction algorithm [71]. Statistical Control of Artifacts in Dense Arrays Studies (SCADS) method uses thresholding methods to detect artifacts [67]. Some of the most common preprocessing steps involve visual inspection to remove bad channels, interpolation of bad channels to minimize loss of information, filtering to remove direct current (DC) shift, slow frequency component, and power noise, and Independent Component Analysis (ICA) to remove eye movement or electrocardiogram (ECG) artifacts. Some preprocessing steps may not directly involve removing artifacts but are performed so that one can handle the huge EEG data better such as dimensionality reduction (reducing the number of electrodes for analysis) and downsampling. Re-referencing of the electrode also plays a major role depending on the underlying neuronal activity in interest. These techniques must be carefully chosen depending on the data set being used and the questions being tackled about the

data set and must be well understood before implementing so as to know the effect these procedures have on the signal. Some of these procedures have been dealt with in detail in this section.

3.1 Independent Component Analysis

Independent component analysis (ICA) is a statistical and computational method used to separate underlying independent components from observed multivariate non-gaussian statistical data. The technique of ICA was first introduced by J. Herault, C. Jutten, and B. Ans in the 1980s, where they introduced the problem of analysis of multivariate data produced by a linear combination of multiple source signals with an analogy to messages carried in the nerve fibers [46][47].

ICA can be considered as a statistical latent variable generative model which describes the observed data as generated by a linear combination of underlying independent latent sources which cannot be directly observed. A basic ICA model is defined as follows: We observe n random variables $x_1, \ldots, x_n$, which are modeled as linear combinations of n random variables $s_1, \ldots, s_n$:

$$x_i = a_{i1}s_1 + a_{i2}s_2 + \ldots + a_{in}s_n, \quad \text{for all } i = 1, \ldots, n \tag{1}$$

where the a_{ij}, i, j = 1, ..., n are some real coefficients. In vector notation it is represented as

$$\mathbf{X} = \mathbf{AS} \tag{2}$$

where $\boldsymbol{X}$ and $\boldsymbol{S}$ are column vectors of observed variable and the latent variables respectively. Here, both, the mixing coefficients a_{ij} and the independent components (ICs) s_i are unknown.

Once $\boldsymbol{A}$ is determined, ICs are calculated using

$$\mathbf{S} = \mathbf{WX} \tag{3}$$

where $\boldsymbol{W}$ is inverse of $\boldsymbol{A}$. The goal of ICA is to estimate the mixing coefficients and determine the ICs with minimum assumptions as possible. A detailed explanation regarding the conditions which have to be satisfied by the random variables such as statistical independence and non-gaussian distribution of ICs and the assumptions made on $\boldsymbol{A}$ to be square and ambiguities of ICA are clearly put forth in [57]. There have been many algorithms developed to compute ICs fast with each having their pros and cons [55][56]. The standard procedure followed by MNE [41] - an open-source Python package for analyzing human neurophysiological data is shown in Fig.3.1. Since ICA is very sensitive to low frequency drift, the observed signals should first be high-pass filtered. Then, the pre-whitened signals are decomposed using PCA. First n principal components (PCs) are then passed to ICA algorithm to obtain ICs. The ICs corresponding to artifacts are removed after careful visual inspection of the ICs. The final step in the procedure is to reconstruct the signals using ICs and residual PCs.

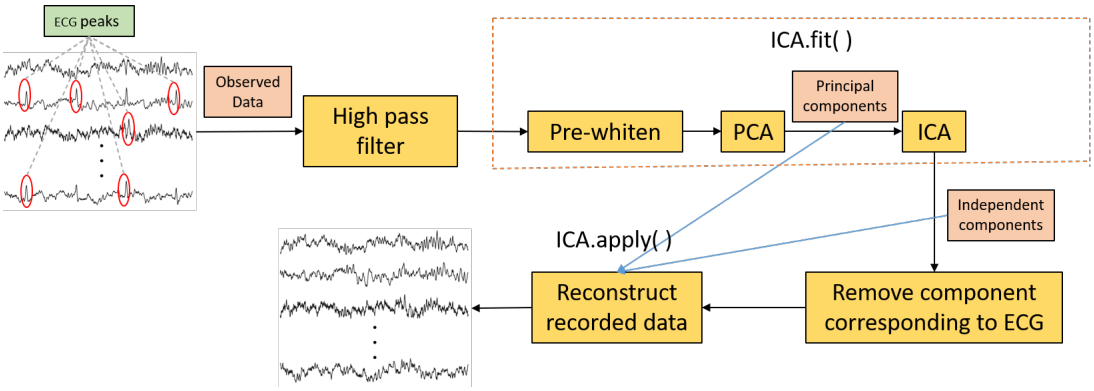

Figure 3.1: ICA procedure followed by MNE

One of the major applications of ICA is found to be in the field of neural signal processing. EEG, as mentioned in the previous section, is a neuro-physiological recording of brain activity consisting of electrical potential recordings from different locations in the brain. Inevitably, EEG is not just a mixture of electrical activities of the brain but also consists of artifacts

such as electrooculogram (EOG), electrocardiogram (ECG), and muscle activity. These potential sources presumably generate the composite signals recorded from the scalp as EEG. Since, neither there is a need to have prior knowledge on the accurate model generating the underlying artifacts nor additional inputs which consists of specified observation intervals of the artifacts, ICA is of the most promising signal processing techniques used for artifact removal in EEG. Fig.3.2 top panel shows an EEG recording typically consisting of ECG components. This can be confirmed by observing the ECG electrode recorded simultaneously shown in the middle panel. The last panel overlays the ECG removed - preprocessed EEG signal over the original unprocessed EEG signal. Even with all the advantages offered by the preprocessing methods such as ICA/PCA, it is suggested that this method be avoided while calculating coherence and phase differences as quantitative EEG (QEEG) features because, the regression and reconstruction effects the raw EEG values and may distort the calculated QEEG features and invalidate them [117].

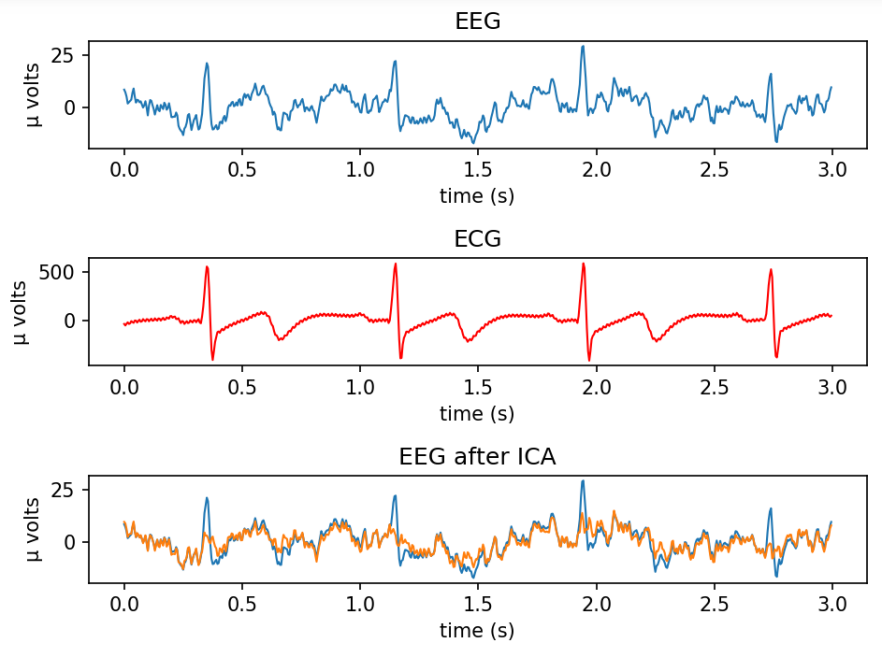

Figure 3.2: Removing ECG component in EEG using ICA

3.2 Thresholding

EEG is a very noisy electrophysiological data that is typically contaminated with artifacts such as eye movements characterized by electrooculogram (EOG), muscle movements characterized by electromyogram (EMG), slow frequency drifts, DC components, and artifacts due to electrode displacement. Most of these artifacts are usually identified by careful visual inspection which is typically a very tedious task.

In case of a bad channel, the recorded signal will be much noisier or appear as a flat line which may arise due to displacement of the electrode from its original position or in case of wet electrodes, due to an increase in the impedance between the electrode-scalp interface which occurs during long term usage of EEG acquisition devices. These channels are usually discarded and interpolated from neighboring electrodes [97]. The relative motion of the electrodes with respect to the scalp also causes offsets in the electrode measurements which can be detected by high amplitude. EMG artifacts manifest as high variance in EEG epochs. There are many algorithms based on ICA and regression which are used to detect and correct artifacts caused by eye or muscle movements. A detailed analysis of these standardized EEG processing pipelines and their effects on EEG signals can be found in [6][83][100][38]. Artifact rejection procedure in epochs used in [83] is shown in Fig.3.3. Amplitude range, variance, and channel deviation are calculated for each epoch and a Z-score of ± 3 is used as a threshold to identify contaminated data.

$$Amplitude\ Range = \langle \max(x_{n_e}) - \min(x_{n_e}) \rangle_N \tag{4}$$

$$Deviation = \langle \langle x_{n_e} \rangle - \langle x_n \rangle \rangle_N \tag{5}$$

$$Variance = \left\langle S^2_{x_{ne}} \right\rangle_N \tag{6}$$

where n, e denotes channel and epoch number respectively. $\langle ... \rangle$ denotes mean operation. Rejection of epoch in one channel results in rejection of the same in all channels. Fig.3.4 shows some of the retained as well as rejected epochs.

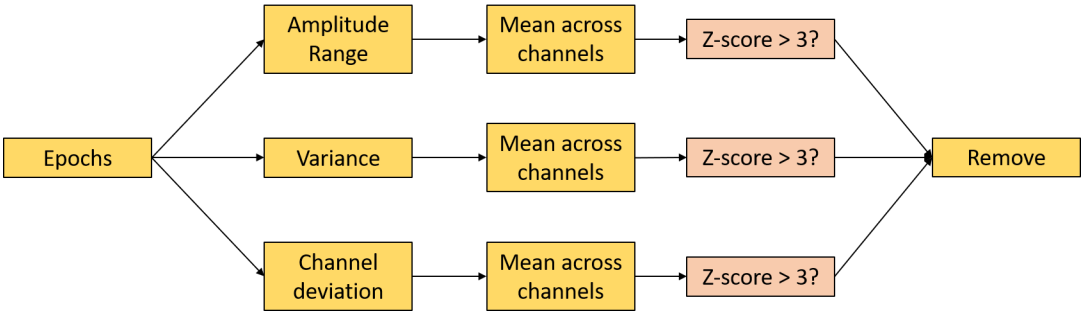

Figure 3.3: EEG epoch rejection procedure

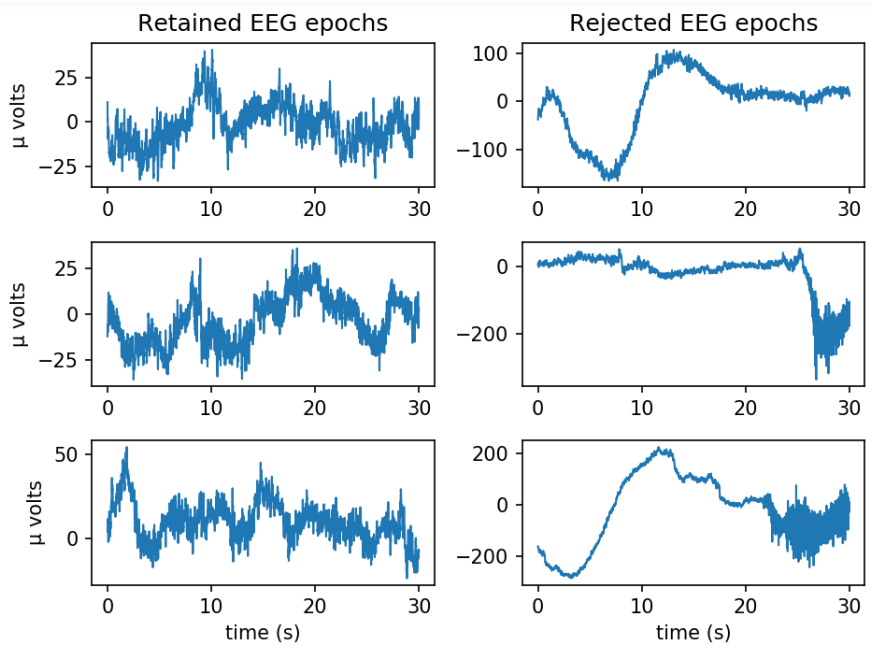

Figure 3.4: Examples of retained and rejected epochs

3.3 Filtering

Filters are one of the most commonly used signal processing tools in neural signal processing. They may be used to remove slow frequency drifts (0-1 Hz), power noise (50 Hz/60 Hz), or to extract various frequency band information such as delta (0.5 - 4Hz), theta (4 - 8Hz), alpha (8 - 12Hz), sigma (13 - 16Hz), beta (16 - 25Hz), and gamma (30 - 35Hz). The filtering process increases the signal-to-noise ratio. As a result, filters play a major role in inferences drawn later from the analysis of the neural signals. The major function of these filters is to attenuate signal components corresponding to certain frequency components which are not of any interest, thereby selectively retaining components corresponding to a particular frequency or range of frequencies in interest present in the signal. As a result of this, usually only the amplitude of the filtered signal is given more importance neglecting the effects the filters have on the phase of the signal. However, along with the modification of the amplitude, filters also change the phase of the signal causing a temporal shift. Linear phase (LP) filters, cause a fixed change in the temporal shift of all the frequencies while, most filters, termed non-LP (NLP) filters, cause a differential time shift as a function of frequency [88]. The induced changes in phase lead to changes in the temporal relationship between oscillations at different frequencies which can lead to misinterpretation of results regarding the underlying process. One should also note here that digital filters usually do not attenuate the undesired spectral components completely, it attenuates to a certain large factor. This is shown in Fig.3.5 where the acquired EEG signal is filtered between 10 - 50Hz. An in depth analysis of digital filter design for electrophysiological data in detailed in [132].

To understand the phase shifting property of filters, consider an input $x(t) = A\cos(\omega_0 t)$ filtered using a filter of frequency response

$$H(j\omega) = M(\omega)e^{j\phi(\omega)} \tag{7}$$

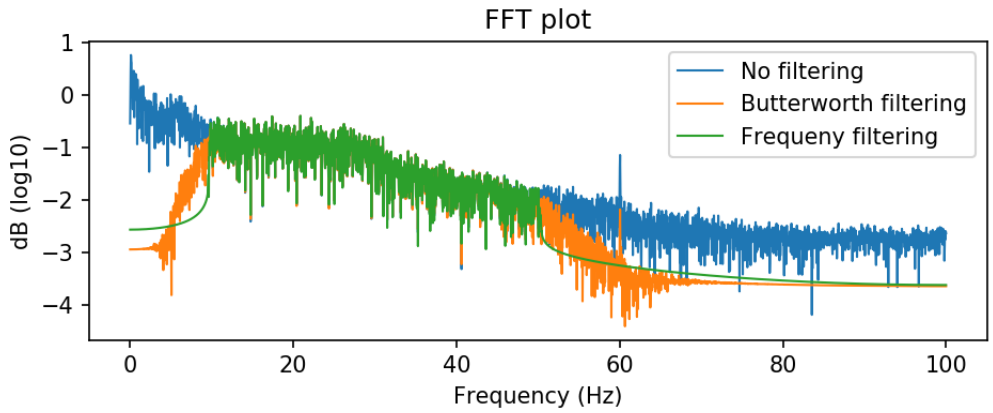

Figure 3.5: Effect of filtering in frequency domain

where $H(j\omega)$ and $\phi(\omega)$ are the magnitude and phase response of the filter respectively. The filtered signal or the response of the filter to the input is given by

$$y(t) = AM\left(\omega_0\right)\cos\left(\omega_0 t + \phi\left(\omega_0\right)\right) \tag{8}$$

which can be rewritten as

$$y(t) = AM\left(\omega_0\right)\cos\left(\omega_0\left(t + \frac{\phi\left(\omega_0\right)}{\omega_0}\right)\right) = AM\left(\omega_0\right)\cos\left(\omega_0\left(t - \tau_p\left(\omega_0\right)\right)\right) \tag{9}$$

From (6) we see that a sinusoidal signal of frequency ω_0 experiences a delay of $\tau_p\left(\omega_0\right)$.

To perform filtering operation in the frequency domain, Fourier Transform (FT) of the signal is obtained [23]. Then it multiplied by a window defining the frequency of interest. Once the desired frequency range is obtained, the signal is reconstructed back to the time domain by taking Inverse Fourier Transform (IFT). Fig.3.6. shows the effect filtering has on the phase of the signal.

The simulated input signal consists of the sum of 2 sine waves one at 5Hz and the other at 15Hz. Then the signal is filtered both in the frequency domain as mentioned in the procedure explained previously and in the time domain using a 6th order butterworth filter

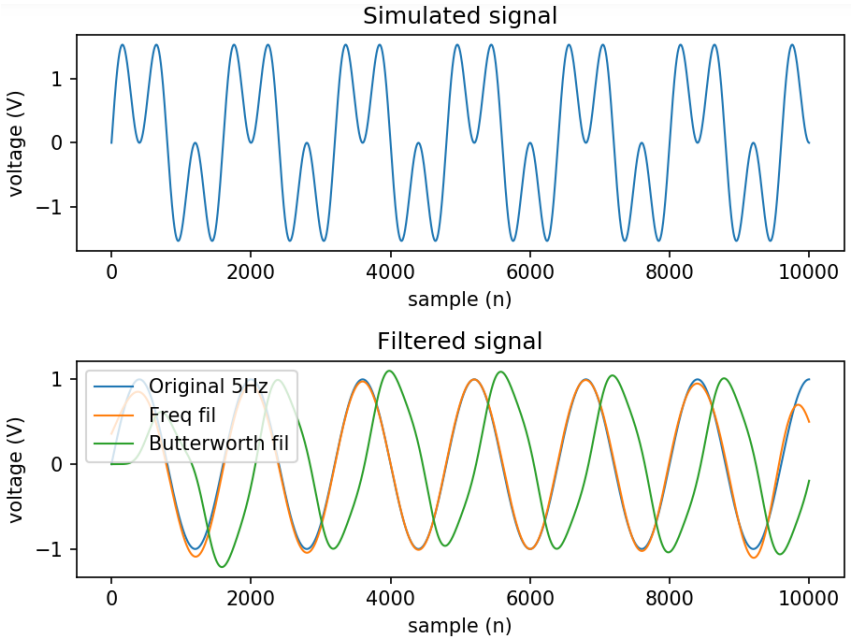

Figure 3.6: Phase shift due to filtering. Upper panel shows a simulated signal of sum of 2 sine waves of 5Hz and 15Hz. Bottom panel shows original 5Hz component present in the signal, results of filtering in frequency domain and using butterworth bandpass filter

independently in the passband of 0.5 - 10Hz. As it can be clearly seen from the Fig.3.6. the resulting filtered signal obtained from filtering in frequency domain overlaps exactly with the 5Hz frequency component present in the input signal whereas, the filtered signal obtained from filtering using butterworth filter has a phase shift.

Similarly, the effect of filtering on EEG is shown in Fig.3.7. Also, filtering in the frequency domain is sometimes more efficient as the convolution steps involved between the input signal and the filter in the time domain transforms to multiplication operation in the frequency domain and we have a faster implementation of FT in the form of Fast-FT (FFT) and Inverse Fast-FT (IFFT). [136] gives a detailed explanation on the consequence of induced phase shifts during filtering process of neural signals.

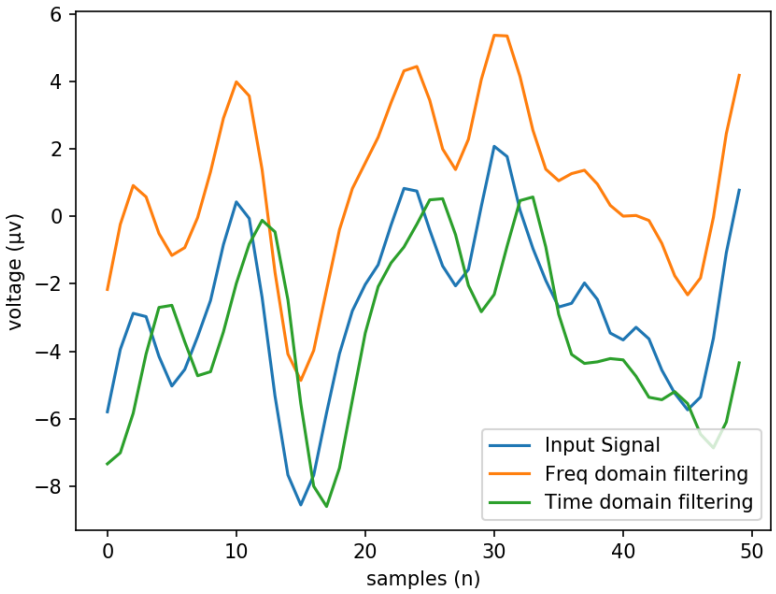

Figure 3.7: EEG filtering

Chapter 4

Feature Extraction

Discriminant models are governed by some specific predefined rules which help in classifying one class from another. Domain knowledge is crucial in the development of such models as one has to decide which QEEG features are most relevant for the detection or classification problem one is working to solve. [98] describes the necessary considerations one has to keep in mind while developing such algorithms. The most commonly used discriminant function to study TBI is the one developed by Brainscope [94]. The other notable QEEG discriminant function used for detection on TBI was developed by Thatcher et. al. [118][113] which consisted of 20 and 16 QEEG features respectively that mostly comprised of the measures discussed in the following subsections. QEEG-based discriminant functions are one of the most promising tools for TBI detection as it takes into account multiple features in the evaluation process thereby improving specificity and sensitivity. However, care must be taken as some studies show a return of prominent QEEG features to normal within few days after concussion [60][29]. Over the past years, there has been the development of advanced signal processing techniques involving complexity measures such as scale-free brain activity [45], graph theory [130], and analysis of causal interactions whose potential in the application of TBI detection and classification has not yet been fully explored [60][82]. Some of the well-

consolidated review work on the use of QEEG for TBI detection and classification can be found in [98][89][128][103][86].

4.1 Spectral Features

The study of frequency domain features is one of the most common forms of EEG analysis. It involves the study of the distribution of the EEG signal over various frequencies which are limited to a narrow frequency range of 0.1 - 100Hz [84][111] which are subdivided into different frequency bands. The boundaries of these sub-bands are slightly blurred and can differ among various researchers resulting in a slight variation of results across studies. the typical range of these frequency sub-bands are delta (0.5 – 4Hz), theta (4 – 8Hz), alpha (8 – 12Hz), alpha1 (8 – 10Hz), alpha2 (10 – 12Hz), sigma (12 - 16Hz), beta1 (12 – 25Hz), beta2 (25 – 35Hz), gamma (35 – 50Hz). A key attribute of frequency/spectral analysis that accounts for its strong applicability is that the spectral information can be derived from very few electrodes making it an easy acquisition set up to record neural activity on the field.

Most of the studies report statistically significant alteration in at least one frequency band in TBI patients compared to control groups. Though the inferences somewhat vary, the general conclusion drawn from these studies shows a decrease of power in higher frequency band such as alpha and an increase of power in lower frequency bands such as delta and theta in TBI subjects [133][119][40][114]. Some studies showed a change in the ratio of theta/alpha power in TBI subjects returned to normal levels after few months suggesting that these measures can be used for longitudinal studies of TBI [20]. However, these studies have to be validated. Some of the Spectral features have been discussed in detail in the subsequent subsection.

4.1.1 Average Power

Many different methods can be used to calculate the power of a signal. One such method is discussed here. To calculate the power in a different frequency band, first power spectral density (PSD) is calculated using Welch's periodogram [131] which averages consecutive Fourier transform (FT) [23] of small windows of the signal, with or without overlapping. Since the spectral content is not stationary in EEG signals, in order to obtain the best estimates of the PSD one has to average the periodograms obtained over short segments of the windows. The optimal window length can be defined to encompass at least two full cycles of the lowest frequency of interest.

Mathematically, the signal is $X(j)$ of length N is divided into smaller segments $X_k(j)$ of length L. FFT of the signal window obtained by multiplying signal $X_k(j)$ with window $W(j)$ of length L is calculated by

$$A_k(n) = \frac{1}{L}\sum_{j=0}^{L-1} X_k(j)W(j)e^{-2kijn/L} \tag{10}$$

K modified periodogram is obtained using

$$I_k\left(f_n\right) = \frac{L}{U}\left|A_k(n)\right|^2 \quad k = 1, 2, \cdots, K, \tag{11}$$

where

$$f_n = \frac{n}{L} \quad n = 0, \cdots, L/2 \tag{12}$$

and

$$U = \frac{1}{L}\sum_{j=0}^{L-1} W^2(j) \tag{13}$$

The spectral estimate is calculated by taking the average of these periodogram

$$\hat{P}(f_n) = \frac{1}{K} \sum_{k=1}^{K} I_k(f_n) \tag{14}$$

More on this is given in [131].

The top panel in Fig.4.1. shows the periodogram obtained for EEG signal shown in Fig.1.2 using Welch's periodogram method. This signal has undergone the preprocessing steps mentioned in the previous section which includes ICA, artifact removal using thresholding, and filtering in the range of 0.5 - 50Hz. For better visualization of the periodogram, PSD is plotted in decibel (dB) in the bottom panel.

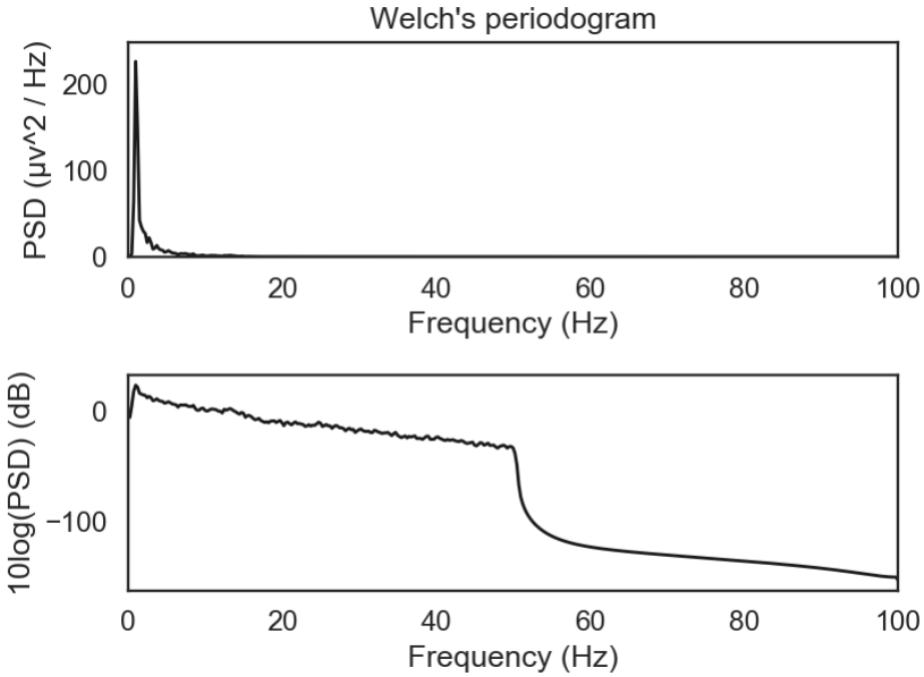

Figure 4.1: Periodogram obtained for signal in Fig.1.2

Note that the PSD value drops off drastically after 50Hz since the signal used to plot these graphs has undergone filtering from 0.5 to 50Hz. The next step in calculating the average power is to define the required frequency bands. Some of the EEG frequency bands mentioned above have been highlighted in Fig.4.2.

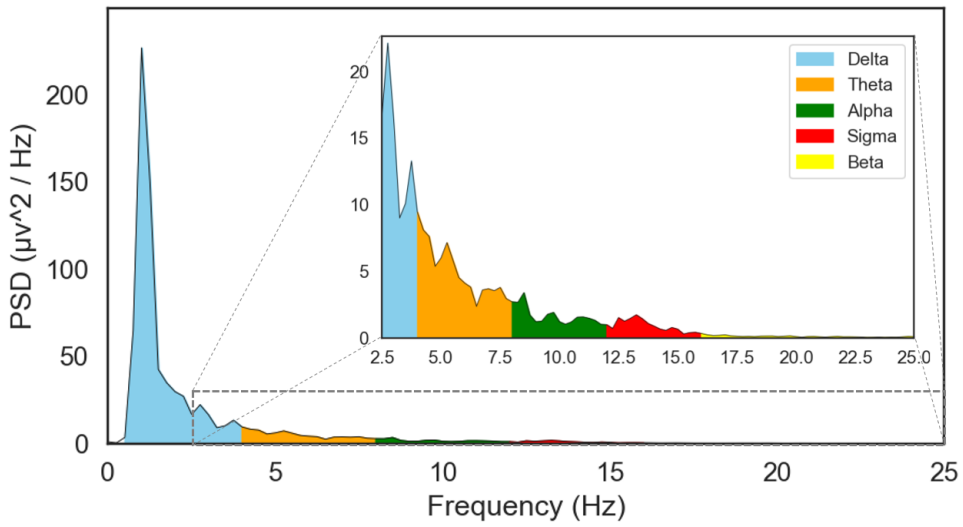

Figure 4.2: EEG Frequency band

The average power in the specific frequency band is then equal to the area of the shaded region corresponding to that frequency band in the Fig.4.2

4.1.2 Relative Power

Relative power in the different frequency bands is given by the ratio of power in the specific frequency band to the total power. Again, total power is defined slightly differently by different researchers. It can be considered as the power of EEG signal in a frequency band corresponding to 0.5 - 50Hz.

$$Relative\ Power = \frac{Power\ in\ frequency\ band}{Total\ Power} \tag{15}$$

4.1.3 Slow:Fast Power Ratio

As mentioned previously some studies showed a change in the ratio of slow power:fast power in TBI subjects [20][78]. Therefore, power ratios of $theta$: $alpha1$, $theta$: $alpha2$, and

$alpha1 : alpha2$ are also considered as a prominent QEEG feature to detect anomalies in EGG in TBI subjects.

4.1.4 Frequency Amplitude Asymmetry

The difference in absolute power between pairs of electrodes is computed to calculate frequency amplitude asymmetry. For inter-hemisphere comparison, it is calculated as

$$Frequency\ Amplitude\ Asymmetry = \frac{Left\ -\ Right}{Left\ +\ Right} \tag{16}$$

and for intra-hemisphere comparisons, it is calculated as

$$Frequency\ Amplitude\ Asymmetry = \frac{Anterior\ -\ Posterior}{Anterior\ +\ Posterior} \tag{17}$$

Frequency amplitude asymmetry was one of the important feature used by Thatcher et. al. to develop a discriminant score for detect TBI [118].

4.1.5 Phase Amplitude Coupling

The interaction of neural oscillations in different frequesncy bands is called cross-frequency coupling (CFC). One of the best example CFC is phase amplitude coupling in which the amplitude of high-frequency oscillations is modulated by the phase of low-frequency rhythms [122][80]. The procedure to obtain the modulation index which quantifies PAC developed by Tort et. al. [122] is shown in Fig.4.3. and Fig.4.4. The raw signal $x(t)$ is filtered in two frequency range of interest, one at a lower frequency (f_P) as $x_P(t)$ and one at a higher frequency (f_a) as $x_a(t)$. Hilbert transform is used to obtain the phase $\phi_{x_P}(t)$ of the signal filtered at f_P and the amplitude envelope of the signal filtered at f_a as $A_{x_a}(t)$. This is shown

in Fig.4.3. The use of Hilbert transform is explained in more detail in the subsection of phase difference. Then the normalized mean amplitude is calculated for each frequency bin as shown in Fig.4.4 denoted as

$$P(j) = \frac{\langle A_{x_a} \rangle_{\phi_{x_P}}(j)}{\sum_{k=1}^{N} \langle A_{x_a} \rangle_{\phi_{x_P}}(k)} \tag{18}$$

where $\langle\ \rangle$ denoted mean operation. The phase-amplitude plot is obtained by plotting P as a function of the phase bins. The modulation index which gives us an idea of PAC is calculated using Kullback–Leibler (KL) distance [72] as the deviation of Phase-Amplitude plot from a uniform distribution.

$$MI = \frac{D_{KL}(P, U)}{\log(N)} \tag{19}$$

where $D_{KL}(P, Q)$ is the KL distance between discrete distribution P and Q is given by

$$D_{KL}(P, Q) = \sum_{j=1}^{N} P(j) \log \left[\frac{P(j)}{Q(j)} \right] \tag{20}$$

4.2 Connectivity Features

Connectivity features are a measure of the relationship between EEG signals recorded simultaneously from different locations on the scalp that requires multichannel EEG recording which involves more complex electrode montages and advanced computational capacity for data analysis as compared to spectral features. Unlike causal relationships which determine the direction of information exchange, connectivity measures are adirectional which means they simply convey if the EEG signals recorded from different locations are in some way correlated or not. Connectivity measures can be broadly divided into three categories: time-

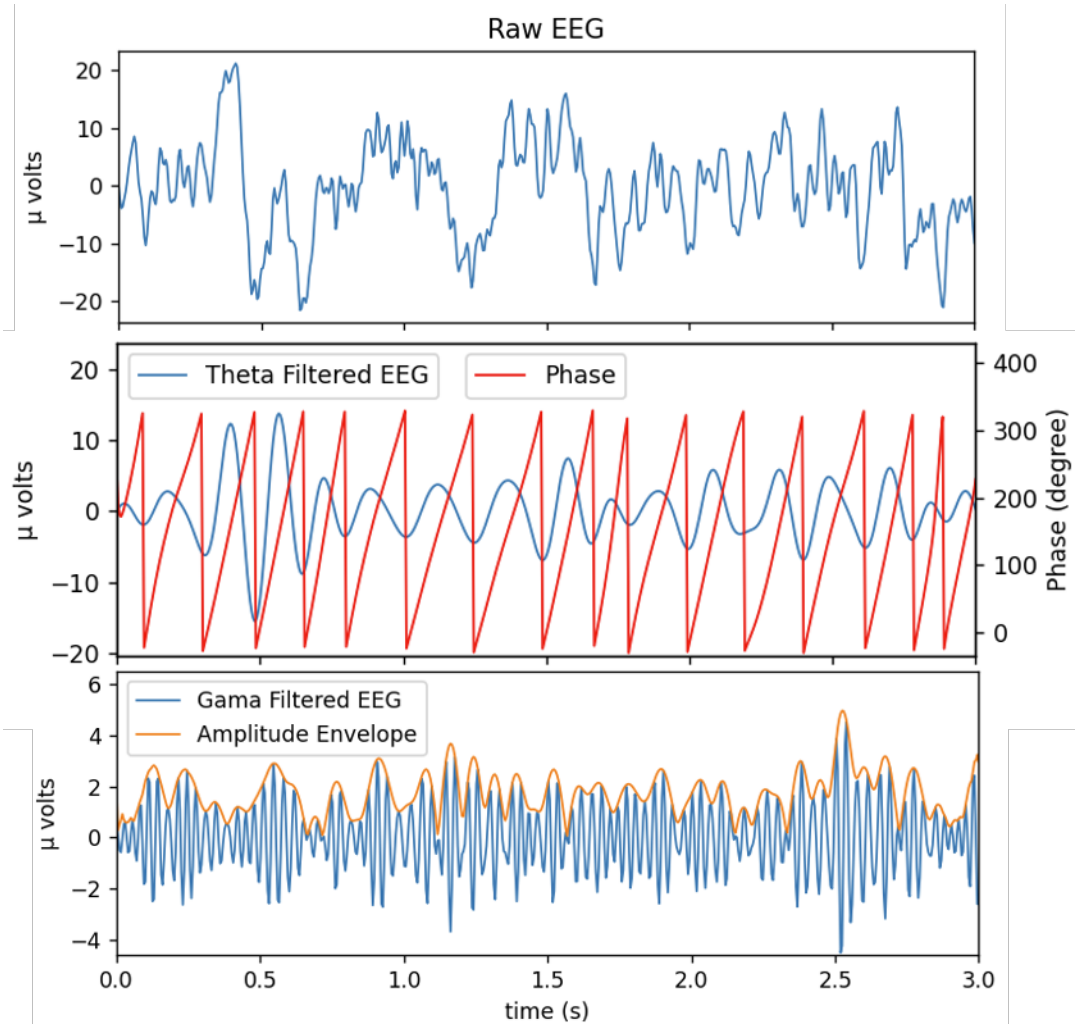

Figure 4.3: PAC procedure

domain measures, frequency-domain measures, and measures calculated from the geometry of embedded data. As the name suggests time-domain and frequency-domain connectivity measures quantify correlation between the original time-domain EEG signal and between spectra or phase of EEG signals respectively. Commonly used frequency domain connectivity features include coherence [85], phase synchronization [2], phase-locking value [73], and phase lag index [109]. Though many studies show a significant change in coherence [78][114][119] and phase-locking value [108] in TBI subjects, sometimes specifically in frontal electrode positions care must be taken regarding the referencing of electrodes before computational analysis can be performed since referencing of EEG electrodes pay a major role in inferences

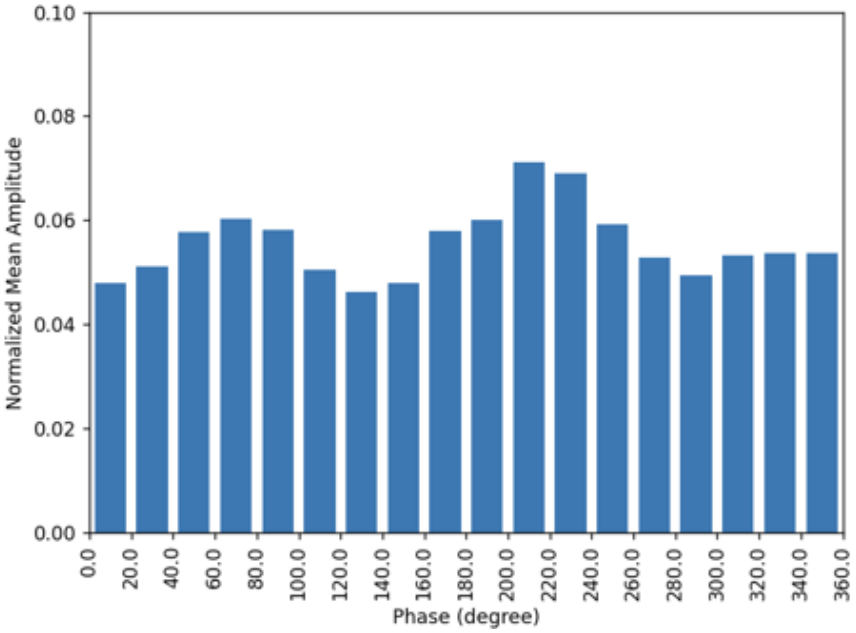

Figure 4.4: Phase Amplitude Plot

of these connectivity measures [3]. The computation of coherence and phase-locking value has been discussed in the following subsections.

4.2.1 Coherence

Coherence [17] is a widely used measure to quantify synchrony between two difference regions of the brain [85]. This gives a means for spatial analysis of the recorded EEG signals. Higher coherence value between pairs of electrodes indicate higher lever of synchronization between the respective regions of the brain. Mathematically, the coherence can be attributed to the frequency domain equivalent to the time domain cross-correlation function. It is a normalized quantity ranging between 0 to 1 given by squared correlation of two spectral density functions over trials

$$\mathrm{Coh}(f,t) = \frac{|\sum_n S_{1n} \cdot S'_{2n}|^2}{\sum_n |S_{1,}|^2 \cdot \sum_n |S_{2n}|^2} \tag{21}$$

where spectral-temporal density function of a signal is expressed as

$$S(f,t) = A(f,t) \cdot e^{i\phi(f,t)} \tag{22}$$

where A and ϕ are the amplitude and phase of the signal at certain frequency f and time t. Many studies have previously used coherence as a measure to detect tbi [119][115][127]. The quality of estimates of coherence highly depend on the number and size of segments correspond to the same process with the same spectral properties used to calculate it [17][85]. Also, as seen in the above equation, coherence is prone to relative change in the amplitude of the signals being considered. To overcome these drawbacks, an alternative measure called phase locking value (PLV) was introduced and is discussed bellow.

4.2.2 Phase difference

Thatcher et. al. [114] reported EEG phase difference between certain pair of electrodes to be comparably more accurate in predicting TBI compared to EEG coherence and relative power. Thornton [119] also reported a change in phase in beta1 and beta2 spectral bands in TBI subjects in a audio memory task paradigm. The phase of a signal is calculated as the angle of its analytical signal which is obtained by taking the Hilbert transform of the signal [21][36][88].

$$\phi(t) = \arg\left[s_{\mathrm{a}}(t)\right] \tag{23}$$

where $s_{\mathrm{a}}(t)$ is the analytical signal of $s(t)$ which is given by

$$s_{\mathrm{a}}(t) = s(t) + j\hat{s}(t) \tag{24}$$

Hilbert transform ($\hat{s}(t)$) of signal $s(t)$ is given by

$$\hat{s}(t) = \left[\frac{1}{\pi t} * s(t)\right] \tag{25}$$

where $*$ conveys convolution operation. Fig.4.5 illustrates phase of a section of filtered EEG signal in theta frequency band. Here the phase of the signal is plotted between $-\pi$ to π radians.

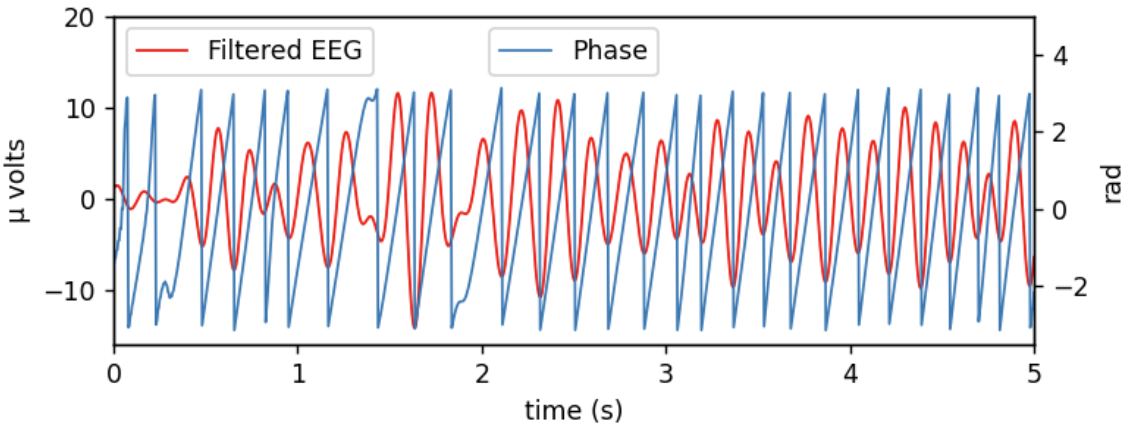

Figure 4.5: Phase of filtered EEG signal

4.2.3 Phase Locking Value

Phase locking value (PLV) as a measure of phase synchrony between brain regions was introduced by Lachaux et. al. [73]. This technique was introduced to over come the shortcoming in the interpretation of phase synchrony by coherence calculation . The relationship between coherence and PLV is well put forth in [15]. Reduction in inter-hemisphere PLV over anterior portions of the brain in delta, theta and beta frequency bands have been reported in TBI cohort previously [108].

$$PLV(f,t) = \left|\frac{1}{N}\sum_{n} e^{i(\phi_{1n}-\phi_{2n})}\right| \tag{26}$$

where ϕ_1 and ϕ_2 and phase of the two EEG signals respectively. An example of PLV calculated for monopolar electrodes of an TBI subject is shown in Fig.4.6.

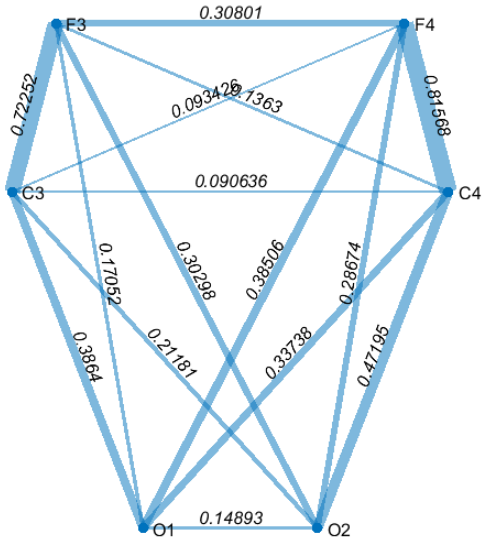

Figure 4.6: Phase Locking Value

4.3 Time domain Features

4.3.1 Hjorth Parameters

Hjorth parameters helps to quantify characteristics of EEG signal in time domain. They may be considered as a bridge between time-domain and frequency-domain characterization as they can also be derived from the first five statistical moments of the power spectrum. Hjorth parameters have shown to be promising measure in the domain of TBI classification [81]. They consist of three measures - activity, mobility and complexity [48].

Activity is measured as the variance of the amplitude of the EEG signal which also characterizes signal power.

$$\text{Activity } (x(t)) = \text{var}(x(t)) \tag{27}$$

Mobility is defines as the square root of the ratio between the variances of the first derivative

and the amplitude of the EEG signal. It is expressed as a ratio per time unit.

$$\text{Mobility}(x(t)) = \sqrt{\frac{\text{var}\left(\frac{dx(t)}{dt}\right)}{\text{var}(x(t))}} \tag{28}$$

Complexity is a dimensionless quantity measures as the ratio between the mobility of the first derivative of the signal and the mobility of the signal itself.

$$\text{Complexity } (x(t)) = \frac{\text{Mobility } \left(\frac{dx(t)}{dt}\right)}{\text{Mobility } (x(t))} \tag{29}$$

where x(t) denotes EEG signal.

4.4 Non-linear Features

4.4.1 Spectral Entropy

The entropy of a signal is insensitive to the order of the measurements in the signal. That is, even if the values are randomly shuffled the same entropy value is obtained when calculated. Entropy quantifies the uniformity of energy distribution in a signal [105]. In particular, spectral entropy measures the uniformity of energy distribution in the frequency-domain. Spectral entropy has been a commonly used quantity in EEG signal processing to measure these irregularities in the signal [59] and previous studies have shown that spectral entropy has been a useful measure in classifying TBI characteristics [127]. Spectral entropy is calculated using the standard entropy formula

$$H(x, sf) = -\sum_{f=0}^{f_s/2} P(f) \log_2[P(f)] \tag{30}$$

where $P(f)$ is normalized PSD and f_s is the sampling frequency.

Chapter 5

Feature Normalization

5.1 Log Transformation

Log transformation is one the most widely used transformation not only in neural signal processing but also in biomedical research in general. QEEG features are generally highly skewed and tailed distributions. Skewness can be considered as a measure of symmetry in a distribution. The importance of approximation to a Gaussian distribution was emphasized by both Dr. E. Roy John and Dr. Frank Duffy in the 1970s and 1980s respectively and will also be discussed in detail in the section of Z-score Normalization. In simple words, log transform compresses a particular range of values and expands other ranges of values depending on the class of log transformation being used. Fig.5.1 shows some of the commonly used log transformations on the QEEG features.

Consider the $log(x)$ curve. Here the smaller values of x are expanded and larger values are compressed. Similarly, $log(x/(1-x))$ transform expands smaller and larger range of values while compressing the mid-range values. Proper care must be taken and the distributions of features being transformed must be kept in mind before choosing the appropriate log

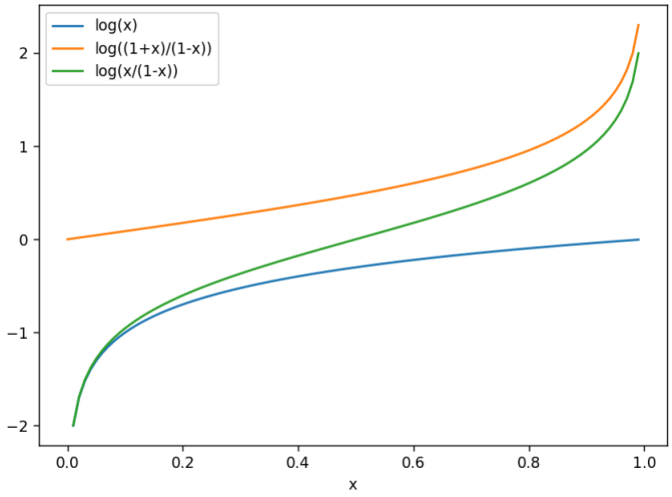

Figure 5.1: Log transformation

transformations. Generally, relative band power R is transformed using $log(R/(1-R))$, magnitude squared coherence C is transformed with $log(C/(1-C))$, amplitude asymmetry X is transformed with $log((2+X)/(2-X))$ and spectral entropy $SpEn$ using $-log(1-SpEn)$. An in-depth analysis of conforming QEEG features to gaussian normality have been discussed in [64][39][117]. The benefits of log transforms can be seen in Fig.5.2 and Fig.5.3. Fig.5.2 and Fig.5.3 are density histogram plots of delta power in F3 electrode and relative alpha power in O2 electrode where they help in reducing the skewness present in the data and conform it to gaussian distribution. However, it is important to keep in mind that once the log transformation is done, the features will no longer have the meaning of the original QEEG features and cannot be directly interpreted as before.

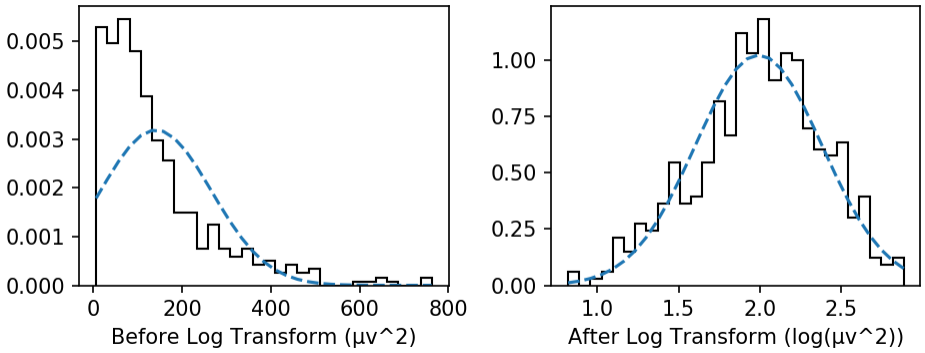

Figure 5.2: Effect of log transformation on features (Delta Power F3-electrode)

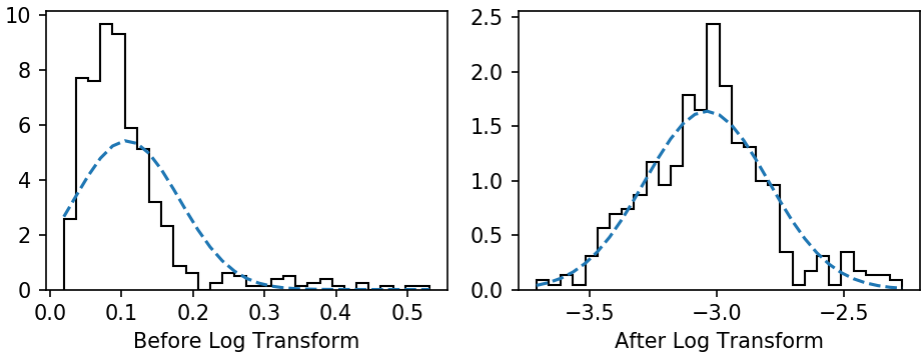

Figure 5.3: Effect of log transformation on features (Relative Alpha Power O2-electrode)

5.2 Age Regression

The age and functional status of the brain are reflected in EEG. The dominant frequency of EEG increases with age and incidents such as brain damage, dysfunction, or deterioration causes frequency slowing in the involved brain regions [65][66][93]. In [65] linear regression equations have been developed to predict the frequency composition of EEG within four frequency bands, for four bilateral regions of the brain, as a function of age which describe the development of the electrical activity of the normal human brain, independent of cultural, ethnic, socio-economic, or sex factors. There are generally 2 different approaches used to minimize the effect of age on QEEG features.

Age stratification involves the grouping of subjects with respect to age and computing mean and standard deviation for each group [115][77]. The grouping of subjects and the number of subjects per age group highly depends on the age of the samples, the relative rate of maturation, and the questions being addressed. A simple method to increase stability and sample size is to use sliding averages for the age stratification. In [116] Thatcher et. al. used one year age groups with sliding averages of approximately .25 years. The best method has to be chosen depending on careful examination of validation at different age groups.

The second method is age regression which was introduced by John et al. Here, the model assumes a linear relationship between the calculated QEEG features and log_{10} of subject's

age expressed in years. The intercept and the coefficients obtained from fitting a straight line to QEEG features with respect to $log_{10}(Subject's\ age)$ are then used to regress the same feature using the below equation.

$$y_i = x_i - \log 10(\text{ SubjectAge }) \cdot m_i \tag{31}$$

where x_i and y_i represent untransformed and transformed variables respectively and m_i represents the age-regression parameter. An example of age regression is shown in Fig.5.4.

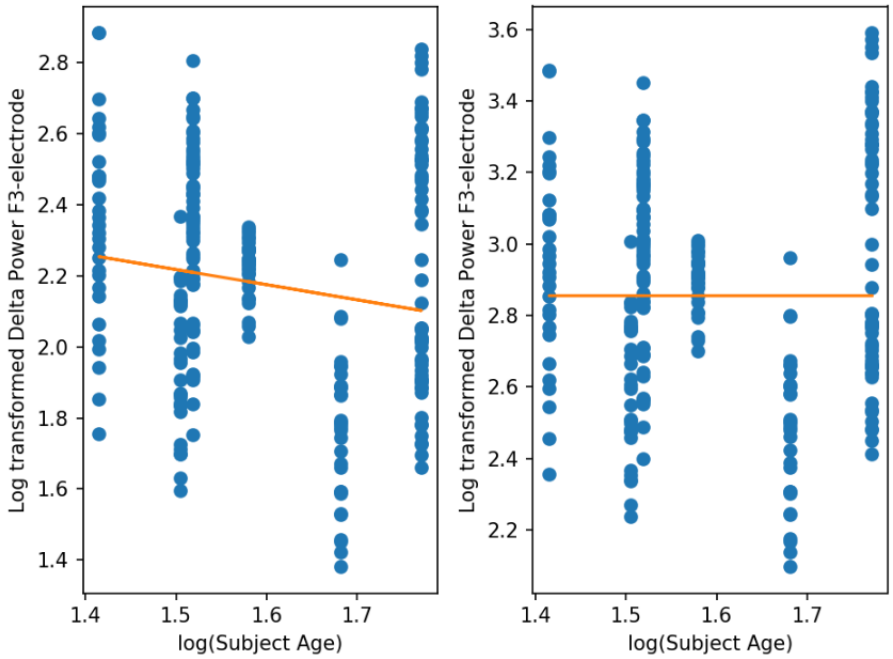

Figure 5.4: Effect of age regression on Log transformed Delta Power F3-electrode

Age regression is performed on the log-transformed QEEG features obtained from the previous step explained in the previous subsection. The left pane shows the QEEG features before age regression is performed. It can be clearly observed that the slope of the best linear fit is not equal to zero indicating the presence of a correlation of this variable with age. As age increases, the delta power decreases as mentioned in [65][66][93]. Once age regression is performed, as shown in the right pane, this correlation is removed. The regression parameters are calculated for norming group (group of all control/normal subjects) which

are then used to age regress TBI subjects [94]. As stated before this model assumes a linear relationship between QEEG feature and age which might not always be true. If the changes in the QEEG features with age is more rapid than a simple linear regression,then one might miss the underlying growth of QEEG feature with age. In those cases, a quadratic or cubic polynomial may account for a better fit which has to be carefully investigated.

5.3 Z-score Standardization

Feature scaling is one of the important steps that will speed up the training process in an ML model. Z-score standardization brings all features to a similar scale of zero mean and unit standard deviation which helps to minimize the cost function of the ML model and find the global minimum faster especially in gradient descent based ML optimization models. For distance-based computations, the distance will be governed solely by the feature having a wide range of values if the difference between the ranges of QEEG features is quite large. This may inadvertently result in prioritizing one feature over the other. Whereas, tree-based ML models are largely insensitive to such standardization. Z-score standardization is performed by subtracting the mean and then dividing by the feature's standard deviation. It is given by

$$z = \frac{x - \mu}{\sigma} \tag{32}$$

where μ is mean of the feature vector given by

$$\mu = \frac{1}{N} \sum_{i=1}^{N} (x_i) \tag{33}$$

and σ is the standard deviation of the feature vector given by

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_i - \mu)^2} \tag{34}$$

where N is the number of samples and x_i denoted samples of the feature vector. Z-score standardization parameters (mean and standard deviation) are calculated for each QEEG feature for the training dataset which are then used to standardize the corresponding features of the test subjects [37]. More on train/test dataset is discussed in Train/Test Split section. Effect of Z-score normalization is shown in Fig.5.5.

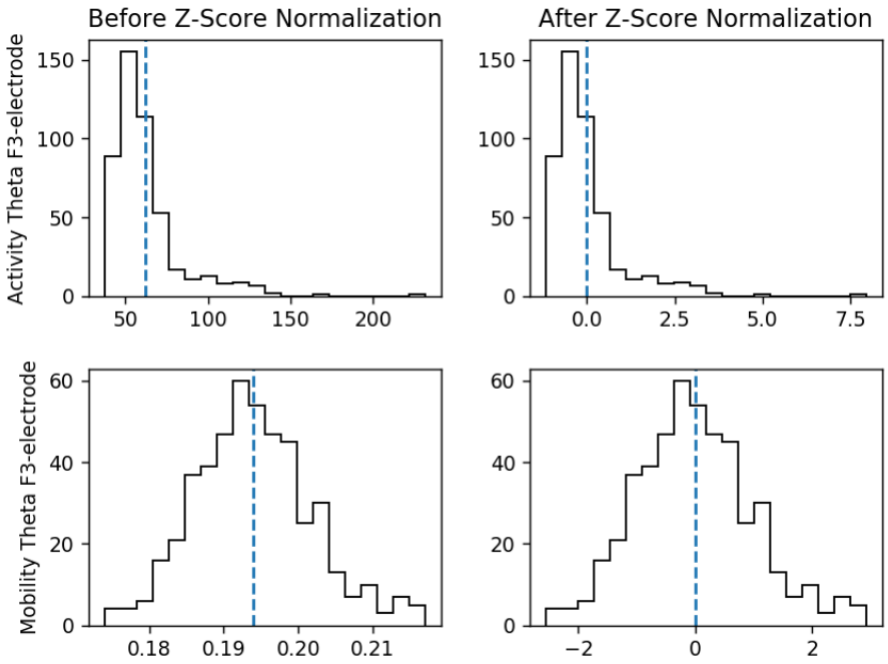

Figure 5.5: Effect of Z-score normalization on log-transformed and age regressed Activity and Mobility theta F3-electrode

All the plots display histograms of the respective features. The features for normalization have been taken from the previous step after log transformation and age regression. The top and bottom rows correspond to activity and mobility feature in theta band for the F3 electrode which is calculated as explained in the feature extraction section. The left column represents features before Z-score normalization and the right column denotes features after Z-score normalization. One has to take note of the range of values taken by both the features. The first QEEG features have a mean and standard deviation of 62.07 and 21.31 respectively. Whereas, the mean and standard deviation of the second QEEG feature is 0.19 and 0.0078. They are drastically different before the standardization step. However,

once the standardization is performed both the features attain zero mean and unit standard deviation and transform to a similar scale giving equal importance to both features in the ML algorithm.

Chapter 6

Feature Selection

Feature selection [18] refers to the process of automatically selecting the most relevant features from the feature pool resulting in the reduction of number of input features used during the development of the ML model [43]. It is an important step in building an efficient ML model mainly for three reasons - to prevent overfitting, to reduce the training time and computational cost of modeling, and to improve the performance of the model, as having models built on irrelevant features decreases its performance. In supervised ML algorithms, statistical-based feature selection is adopted in which the relationship between different features is evaluated with respect to the target variable using cross validation on the training data. Based on this, the features that have the strongest relationship to the target variable are chosen that are further evaluated based on the performance on the hold out set. Selection procedure is not performed on the entire dataset to prevent data snooping. Care must be taken to differentiate feature selection from dimensionality reduction technique. Dimensionality reduction technique reduces the dimensionality of the data by projecting it into new space creating new input features, unlike feature selection methods which only removes less relevant features.

Feature selection procedure may be broadly divided into wrapper, filter, and embedded methods. Filter methods rank the features by scores which are obtained based on statistical quantities depending on their correlation to the target variable and selects the features according to the ranking order. This method is independent of any ML algorithm. Chi squared test, pearson's correlation [44], mutual information [4], and Analysis of variance (ANOVA) are some of the examples of filter method. Wrapper methods [70] are computationally intensive methods which evaluates the performance of a particular ML model on a subset of feature on metrics such as accuracy. Based on the inference drawn from previous model, decision is taken to add or remove a particular feature. This procedure is repeated until the required number of features are obtained. As a result wrapper methods are not independent of ML algorithm. Forward feature selection, backward feature elimination, and recursive feature elimination (RFE) are some of the examples of wrapper methods. Embedded feature selection procedure incorporates feature selection as a part of training process of the model. Hence, it is quicker than the wrapper methods and is not independent of ML algorithm. Regularization methods are the most common type of embedded feature selection procedure which introduces additional constraints that lower model complexity and penalize a feature given a coefficient threshold thereby preventing overfitting. Decision trees [75], least absolute shrinkage and selection operator (LASSO) [121], and ridge regression [52] are some of the examples of embedded feature selection methods.

Fig.6.1. shows the relationship between number of features and the cross-validation accuracy (5 fold) obtained on N2 sleep stage of human dataset using RFE method with random forest as the base estimator. The decision on the optimal number of features for a particular dataset is still an ongoing research [53][32]. [61] shows that the optimal number of features for a pool of uncorrelated features is $N - 1$ whereas for highly correlated features it is $\sqrt{N}$ where N in the total number of features.

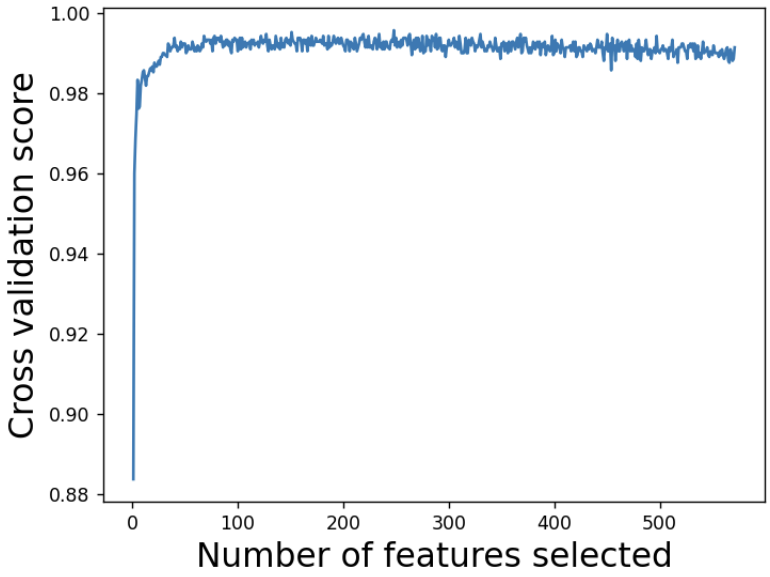

Figure 6.1: Relationship between number of features and CV accuracy

Chapter 7

Machine Learning Approaches

7.1 Train/Test Strategy

Training ML models on the entire dataset results in data snooping and the model outperforming on the test set. If the data were already seen by the ML model, then it overfits and would result in a much higher accuracy. This would also lead to a failure in classifying new, previously unseen data [37]. There are primarily two ways one could test their ML algorithms in these scenarios. They are k-fold Cross Validation and individual validation. The significance of each method is discussed below. In both the data arrangements, it is made sure that the same number of epochs are extracted from each subject to keep it consistent across subjects.

7.1.1 k-fold Cross Validation

In this data arrangement, the features or the raw EEG epochs are shuffled and equally partitioned into k folds. In each iteration one fold is selected as validation data, and the

remaining (k-1) groups are considered as training data [69][62]. Each iteration creates a model with corresponding train/validation dataset and finally the mean accuracy of the models is calculated.

$$\text{accuracy}_{\text{cv}} = \sum_{\text{i=1}}^{\text{k}} \frac{\text{accuracy}_{\text{i}}}{\text{k}} \tag{35}$$

where i denotes iteration number. As a result, the training dataset comprises of approximately $x\%$ of the data or features from each subject, and the rest of the $(100 - x)\%$ data or features from each subject is used as the validation set depending on the number of folds (k) chosen. This is portrayed in Fig.7.1. Each box in the figure corresponds to an epoch. It should be noted here that this is not suitable for highly unbalanced datasets.

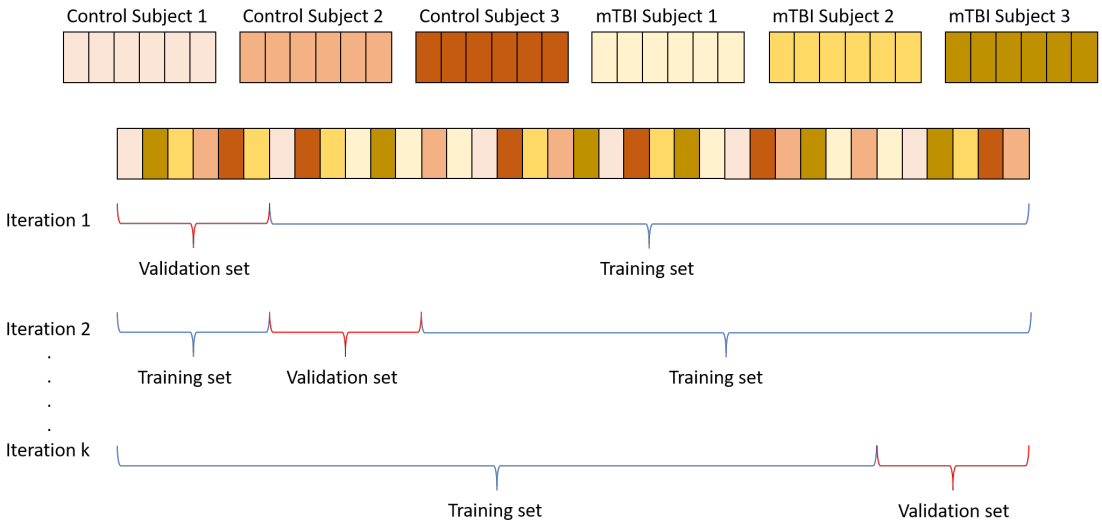

Figure 7.1: k-Fold Cross Validation

Since, a part of the data, specific to an individual is already seen by the algorithm while training, this data arrangement helps to evaluate the power of pattern recognition of the algorithms [8]. As the data from each individual is not separated in train/validation sets, the importance of subject specific differences are not taken into account while validating the algorithm. This type of data arrangement enables one to build models which can be used for daily monitoring of patients since it relies on previous data from the same subject rather than detection problems which involves validation on a completely new subject.

7.1.2 Independent-Validation

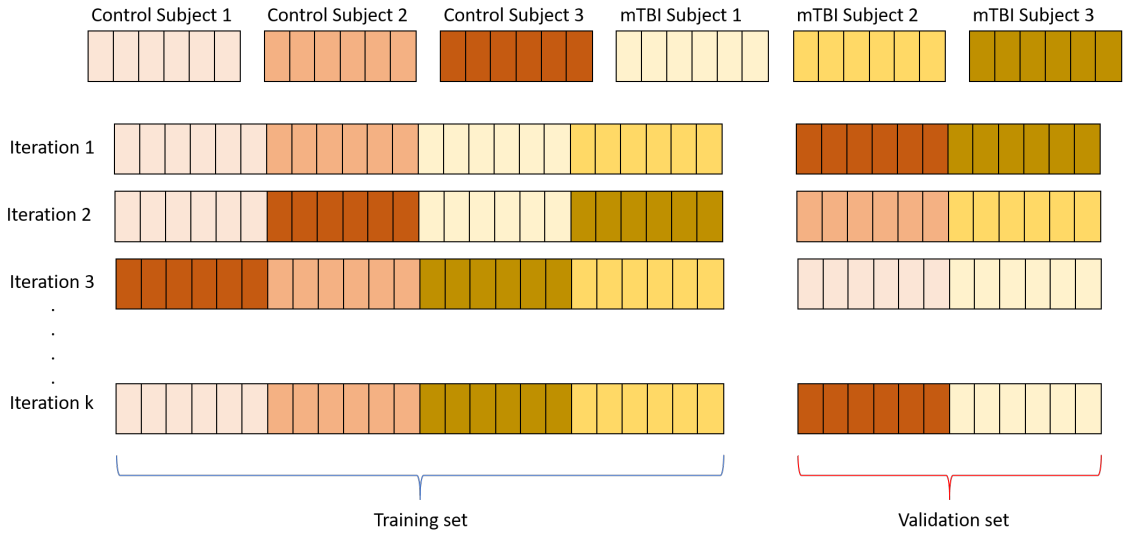

Figure 7.2: Individual validation

In this is data arrangement, one subject from each class - control and TBI are set aside as the validation set and the ML model is trained on rest of the dataset. By having data from both classes in the validation set it is made sure that the composition of the validation set reflects the composition of the training set. This is carried out for all possibles combinations of training and validation sets depending on the number of subjects used in the dataset. As in k-fold cross validation, the final accuracy of the model is considered as the mean accuracy of the models obtained from each iteration. Since the data from an individual subject do not appear in both training and validation set, this data arrangement is known as individual validation (IV). The IV data arrangement gives significant importance to differences in EEG from different subjects. This arrangement helps study the generality of the trained model as the prediction of class is made on subjects whose data is completely new and is not seen by the algorithm beforehand. This arrangement mostly reflects a more practical scenario of TBI detection rather than TBI monitoring as in case of k-fold cross validation arrangement.

7.2 Rule Based ML

Rule-based classical ML models suit best for small to medium structured/tabular data. In the following sub-section we will discuss the working of a number of classical rule-based ML algorithms.

7.2.1 Decision Tree

Decision Trees (DTs) are a non-parametric supervised learning method. Many decision tree algorithms have been developed and improved upon over many years [76], all of which work on broadly the same principle. Some of the well known DT algorithms are Iterative Dichotomiser 3 (ID3) [95], C4.5 [96], and Classification And Regression Tree (CART) [13][75]. DT is built as a tree of nodes in a top-down induction manner successively splitting the root node consisting the source data. This splitting is based on certain conditions with respect to a particular feature. Different algorithms use different criteria to select the best feature for the split at each node. Some of the common metrics used are Gini impurity, information gain, and variance reduction. Used by CART, Gini impurity is calculated as

$$\mathrm{I}_G(p) = \sum_{i=1}^{J} \left(p_i \sum_{k \neq i} p_k \right) \tag{36}$$

where p_i is the probability of an instance with label i, and J is the number of classes. It is defined as likelihood of incorrectly classifying a randomly chosen instance if it were randomly labeled according to the distribution of labels in the subset. It has a lower bound of zero which is obtained when a node consists of data from a single class. During the training process, the best split is chosen by maximizing the Gini gain which is given by the difference between the weighted impurities of the branches from the original impurity. This process is repeated recursively until all data in a node has the same label or when splitting no longer

adds value to the predictions. DTs are easy to interpret but are prone to overfit. Fig.7.3 shows a simple decision dtree built on wake sleep stage of mouse data with two normalized features. Each node displays the splitting condition on the feature, gini impurity, number of samples in the node and the class assigned to the node.

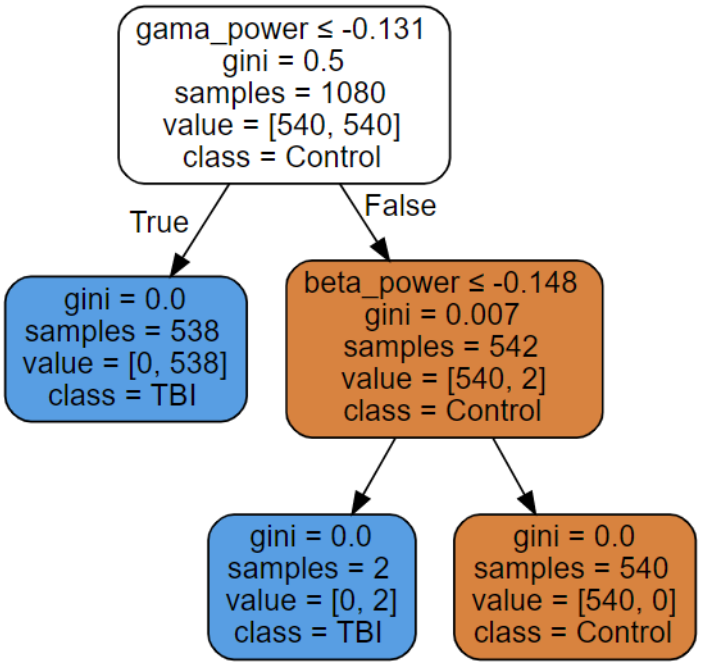

Figure 7.3: Decision Tree

7.2.2 Random forest

Random forests (RFs) [49][12] are a supervised ML algorithm that builds an ensemble of decision tree classifiers which are fit on different sub-samples of the dataset and features with replacement. This is known as bagging [11][50]. This method helps in reducing the variance of the decision trees and avoids overfitting of the data by introducing randomization. The output of the random forest algorithm is however the averaged prediction of individual classifiers. If feature bagging is not performed, the trees developed may be highly correlated if a subset of features are very strong predictors of the target class. In this case, the same

features will be selected in developing different trees making them correlated [51]. Fig.7.4 shows the overlapping RF decision boundaries of 30 different decision trees on wake sleep stage of mouse data with two normalized features identical to the one used to obtained decision tree.

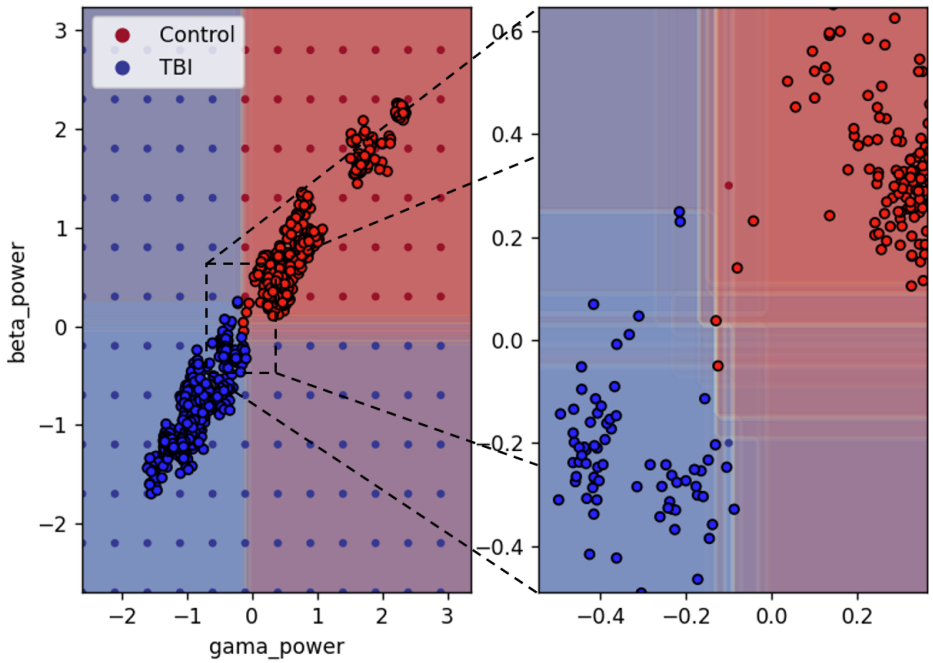

Figure 7.4: Random Forest

7.2.3 Support Vector Machine

Support vector machines (SVM)[24][9] are supervised learning models used for classification and regression. The main objective of SVM is to find a hyper-plane in an N-dimensional space that gives the widest separation between different classes and distinctly classifies the data points. The dimension of the hyper-plane depends on the number of features being used. The distance from the hyper-plane to the nearest data points determines the margin. If the dataset is linearly separable, the hyper-plane is chosen such that it maximizes the margin between classes. This is referred to as hard margin. This boundary is usually specified or determined by a smaller subset of data points known as support vectors as in a vector space

a point is considered to be a vector from the origin and that point. The decision rule is given by

$$\bar{\omega} \cdot \bar{u} + b \geq 0 \tag{37}$$

where $\bar{\omega}$ is the weight vector which is a vector perpendicular to the hyper-plane, b is the intercept term and $\bar{u}$ is the unknown data point. Introducing additional constraints and variables to facilitate calculation of $\bar{\omega}$ and b conveniently we get

$$y_i(\bar{x}_i \cdot \bar{\omega} + b) - 1 = 0 \tag{38}$$

where $\bar{x}_i$ is the data point and y_i is equal to +1 for positive samples and -1 for negative samples respectively thus satisfying the constraints of decision rule being equal to +1 for positive sample and -1 for negative samples. The geometric margin or the width of the decision boundary is given by $2/||\bar{w}||$ which is to be maximised. Hence, the formulation of SVM becomes a minimization problem where, to find $\bar{\omega}$ and b, $\frac{1}{2}\bar{w}^T\bar{\omega}$ is minimized with the constraints specified in the above equation. Lagrange multipliers are used to find extremum of the function subject to equality constraints [16][74]

$$L = \frac{1}{2}\|\bar{\omega}\|^2 - \sum \alpha_i \left[y_i \left(\bar{w} \cdot \bar{x}_i + b\right) - 1\right] \tag{39}$$

where α_i are the Lagrange multipliers. The final decision rule is of the form

$$\sum \alpha_i y_i \bar{x}_i \cdot \bar{u} + b \geq 0 \tag{40}$$

The proof for the decision rule is clearly explained in [124][104]. When the data points are not linearly separable in a particular space they are either transformed into other space using different kernels in which they are linearly separable or soft margin is used [24]. Some of the commonly used kernels are linear, polynomial, and Radial Basis Function (RBF). In case of

soft margin the cost function to be minimized changes to

$$E(\bar{w}, b) = \left[\frac{1}{n}\sum_{i=1}^{n}\max\left(0, 1 - y_i\left(\bar{w}\cdot\bar{x}_i + b\right)\right)\right] + \lambda\|\bar{w}\|^2 \tag{41}$$

where the fist term on the right side of the equation corresponds to the hinge loss and the second term is the regularization term [91]. Fig.7.5 shows the SVM decision boundary and the support vectors obtained using linear kernel.

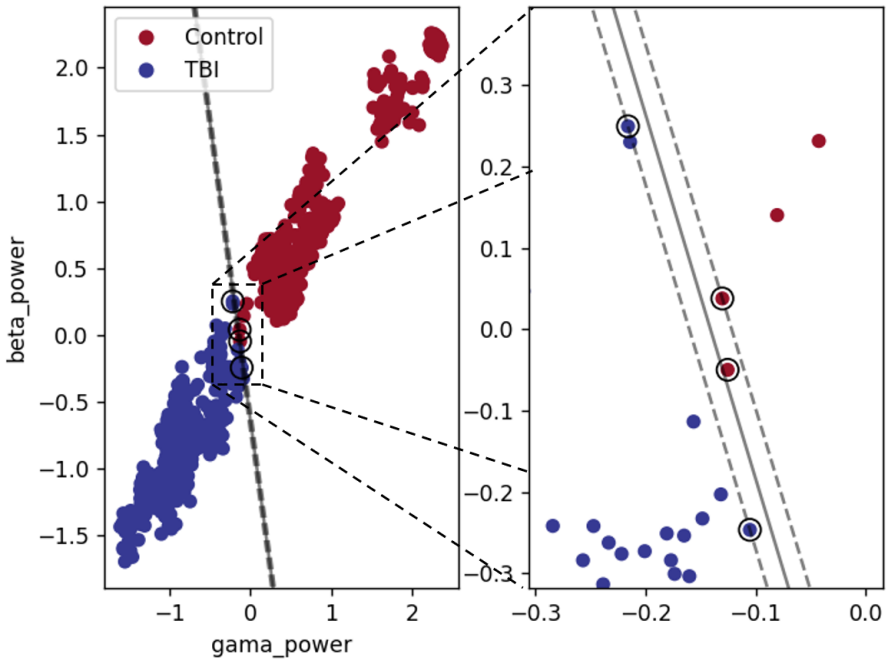

Figure 7.5: Support Vector Machine

7.2.4 k Nearest Neighbor

First developed by Evelyn Fix and Joseph Hodges [33], k nearest neighbor (kNN) is a supervised, instance based learning algorithm. Major assumption made in kNN is that, the data points from same classes exist in close proximity with each other. Instead of building a model, kNN simply stores the instances of the training dataset during the training phase. In the classification phase, the label of the new data point is inferred from a simple majority

vote of the 'k' (a predefined number) of training samples closest to the unknown data point. For example, if k = 1, then the label of the unknown data point is assigned based on the closest training sample to it. In 1967, Thomas Cover and Peter Hart showed, as the size of training dataset approaches infinity, the error of one nearest neighbor classifier is upper bounded by twice the Bayes error rate [26]. Euclidean distance is the most commonly used distance metric. Euclidean distance between two points A and B with coordinates (x_1, y_1) and (x_2, y_2) respectively in the Cartesian plane is calculated as

$$d(A, B) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \tag{42}$$

Since the algorithm is highly dependent on the distance between data points, kNN is highly sensitive to local structure of the dataset. Any scaling operations performed on the data results in a change of the distance between the data points and thereby impacting the assignment of the label. The optimal value of 'k' is highly dependent on the training data, however a larger value of 'k' is chosen to prevent overfitting of the algorithm on the training data. Fig.7.6 shows the decision boundary obtained for two different values of 'k', 5 and 11. As seen in the figure, having a lower value of 'k' bends the decision boundary around the training samples. kNN algorithm which use brute force method to find the k nearest neighbors becomes slower as the size of the samples increases which is one of its major drawbacks. This problem is addressed to some extent by more computationally efficient methods such as k-dimensional tree (k-D tree) [5] or ball tree [87] methods. If the dataset is highly skewed, then the class with more samples dominate the prediction of label of new data points. This is overcome by using weighted nearest neighbor method in which the class of the k nearest data points are multiplied by a weight proportional to the inverse of the distance between itself and the unknown data point thereby giving more importance to the samples which are closer.

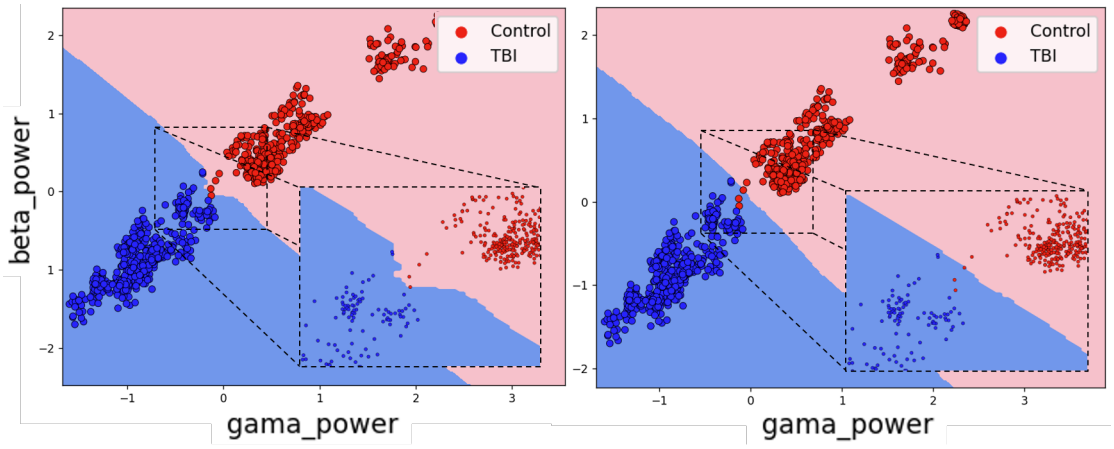

Figure 7.6: K Nearest Neighbor (Left: K = 5, Right: K = 11)

7.2.5 Extreme Gradient Boosting

Also known as XGBoost (XGB) [19], it is an ensemble learning method which implements the gradient boosting decision tree algorithm enabling faster learning through parallel and distributed computing [25]. Boosting is a sequential technique in which the result of the model is the weighted average of previous models. The current decision stump is built by weighing missclassified points form previous stumps higher thereby updating the residual error in each iteration. Since gradient descent algorithm is used to minimize the loss, is it called gradient boosting. XGBoost uses regularization parameters to prevent the model from overfitting. XGBoost betters its performance through system optimization techniques such as parallelization, tree pruning, and hardware optimization, algorithmic enhancement techniques such as regularization, sparsity awareness, and weighted quantile sketch.

Chapter 8

Evaluation Metrics

An accurate interpretation of the results and a careful evaluation of the model used to obtain them are some of the most important steps in any data analytic/pattern recognition problem. Evaluation metrics, also known as performance metrics are fundamental in assessing the quality of the learned methods. There have been numerous metrics defined for this purpose that convey different information regarding the model, enabling one to interpret the performance of the model from different standpoints. Some of the most commonly used metrics are accuracy, sensitivity, specificity, precision, F1-score, receiver operating characteristic curve (ROC), and area under the ROC (AUC). The differences between these metrics become more evident when multi-class problems and problems which deal with imbalances datasets are considered. Some of the frequently used metrics in the medical field have been mentioned in detail in [106]. A number of previous studies have established the importance of understanding these metrics before using them in assessing the ML models [92][112][107]. [123] clearly explains the difference between sensitivity, precision, and predictive values. Some of these metrics have been discussed in more detain in the following subsections. Most of these metrics are derived from confusion matrix. Confusion matrix for a binary classifier is shown in Table 8.1.

		Actual values		
		Positive	Negative	
Predicted values	Positive	True Positive	False Positive	Row entries for determining positive predictive value
	Negative	False Negative	True Negative	Row entries for determining negative predictive value
		Column entries for determining sensitivity	Column entries for determining specificity	

Table 8.1: Confusion Matrix

The confusion matrix summarizes the predictions into four cells which may be expressed using raw counts of the number of times each predicted label is associated with each real class, or may be expressed in relative terms. True positive and true negative refers to the number of correctly predicted positive and negative labels respectively represented in green in Table 8.1. Similarly, false positive and false negative refers to the number of incorrectly predicted positive and negative labels respectively represented in red in Table 8.1. All four cells add up to the total number of data points.

8.1 Accuracy

Accuracy is the ratio of total number of correct predictions to the total number of cases.

$$\text{Accuracy} = \frac{\text{True Positive} + \text{True Negative}}{\text{True Positive} + \text{True Negative} + \text{False Positive} + \text{False Negative}} \tag{43}$$

8.2 Sensitivity

Sensitivity (Sen) also known as recall or true positive rate is the ratio of true positive to the sum of true positive and false negative.

$$\text{Sensitivity} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \tag{44}$$

It conveys the ability of the ML model to detect actual positive cases relative to the known reference standard. However, this should not be mistaken with positive predictive value/precision as sensitivity only takes into account the actual positive cases and assess how well the model performs to detect the positive cases out of all actual positives.

8.3 Specificity

Specificity (Spec) is the ratio of true negative to the sum of true negative and false positive.

$$\text{Specificity} = \frac{\text{True Negative}}{\text{True Negative} + \text{False Positive}} \tag{45}$$

It conveys the ability of the ML model to detect actual negative relative to the known reference standard. However, this should not be mistaken with negative predictive value as specificity only takes into account the actual negative cases and assess how well the model performs to detect the negative cases out of all actual positives. The main focus of sensitivity and specificity is on the accuracy of the ML model relative to the actual reference and is not concerned with the actual true labels of the subjects. As a result if there is an error in the actual label, it is not detected when using sensitivity or specificity. This concept is well explained in [123].

8.4 AUC score

ROC is a visualization tool that helps to compare the performance of different classifiers [31]. It is a two dimensional graph in which true positive rate is plotted against false positive rate depicting the relative trade-off between the two.

$$\text{False Positive Rate} = \frac{\text{False Positive}}{\text{False Positive} + \text{True Negative}} \tag{46}$$

A classifier outputs a score which defines the degree to which an instance belongs to a particular class. A threshold to this score is used to produce a discrete classifier with a rule of assigning positive label if the score is above the set threshold or a negative label if the score is below the set threshold. Different threshold value produces different points on the ROC curve for the same classifier. The optimal threshold corresponds to the threshold obtained at the point closest to the top left corner of ROC space. The diagonal representing y = x in the ROC curve represents the strategy of randomly guessing a class. An illustration of ROC curve is shown in Fig.8.1.

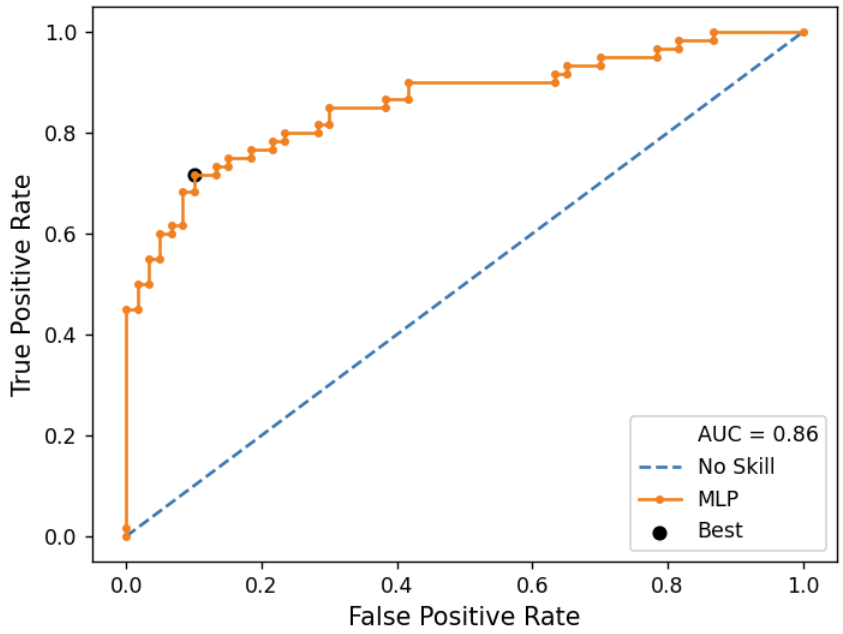

Figure 8.1: ROC curve

One of the most important characteristic of ROC curves is that it is insensitive to changes in class distribution. Since the ROC curve is a two-dimensional depiction of classifier performance it is easier to compare the performance of the classifiers using area under ROC (AUC) which is a scalar value instead [10]. AUC ranges between 0 to 1 since it is defined as the area under the ROC curve which is a portion of the area of the unit square.

Chapter 9

Pipeline and Tools Used

Though the procedure used to analyze the mouse and the human data are similar there very subtle difference due to the difference in the data format of the two species. Since mouse data are single electrode recording connectivity features for mouse data are not calculated and as human cohort belong to a wide range of age demographics, age regression is performed for human data. Also the influence of ECG in human EEG recordings is significant, human data undergoes ICA. The steps performed on data of both species have been explained in the previous sections in detail. Fig.9.1 shows the pipeline followed in both cases

Python 3.8 along with Spyder IDE and machine learning tool: scikit-learn [90] is used to implement and test the algorithms. EEG analysis and visualization tool MNE [42] is used for handling the raw data. Tensorpac [22] is used to calculate PAC. Matplotlib [54] and Seaborn [129] are used to plot figures.

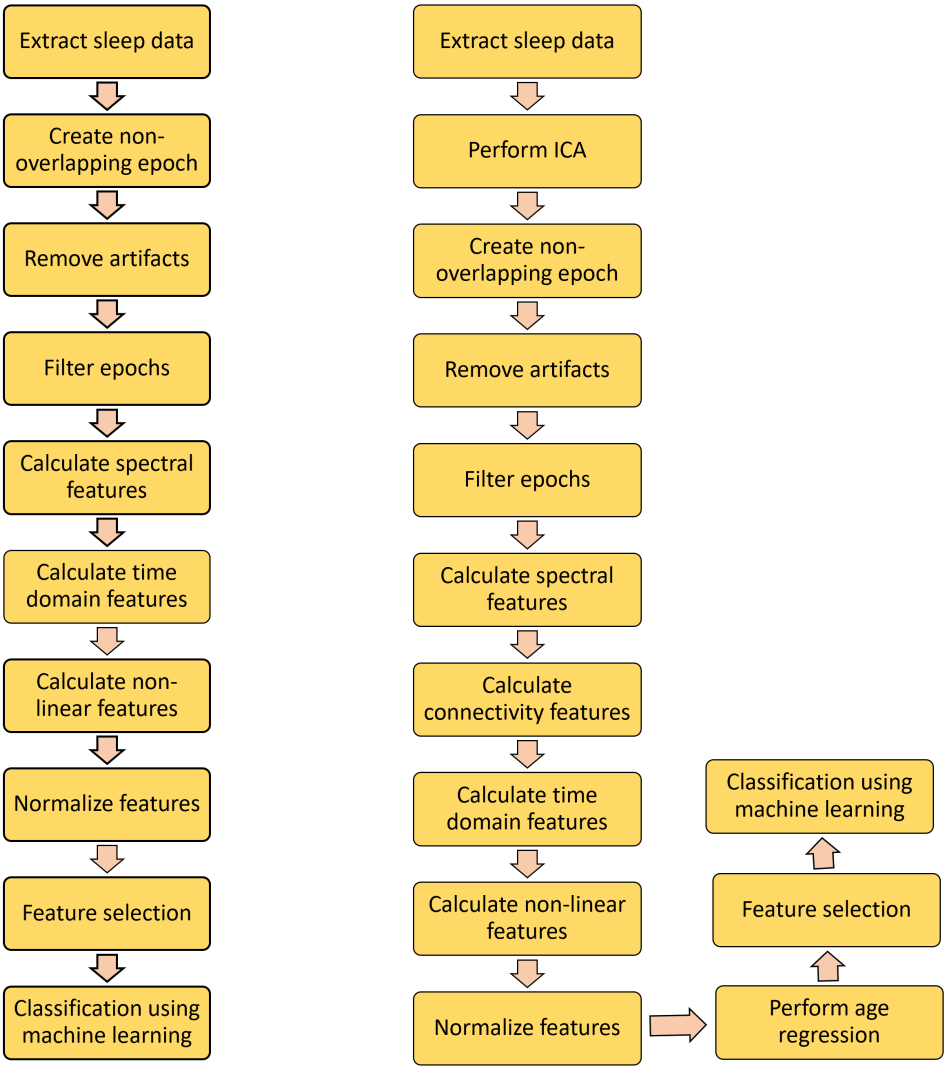

Figure 9.1: Procedure used to analyze mouse data (left) and human data (right)

Chapter 10

Results and Discussion

This section discusses the results obtained using the above pipeline. Table 10.1 and 10.2 shows the results obtained for the mouse dataset. Table 10.1 shows the accuracy obtained for different algorithms and sleep stages in mouse dataset for cross-validation (CV) and IV data arrangement whereas Table 10.2 presents different evaluation metrics for the same dataset in IV data arrangement.

Sleep Stage	DT	RF	kNN			MLP	SVM	XGB
			k = 5	k = 11	k = 19			
	CV							
Wake	99.54	99.63	99.54	99.63	99.63	99.44	98.8	99.63
NR	99.72	99.81	99.81	99.72	99.81	99.07	99.63	99.91
REM	100	100	100	94.55	96.36	85.27	97.27	100
	IV							
Wake	**98.26**	**99.16**	**96.76**	**97.89**	**97.94**	**98.53**	**93.23**	**98.65**
NR	97.54	96.44	91.27	92.03	92.85	95.02	90.23	98.38
REM	95.67	94.48	88.58	91.63	93.83	89.69	87.14	95.27

Table 10.1: Accuracy (%) obtained for mouse data

Table 10.3 and 10.4 shows the results obtained for the human dataset. Table 10.3 shows the

Sleep Stage	Metric	IV							
		DT	RF	kNN			MLP	SVM	XGB
				k = 5	k = 11	k = 19			
W	AUC	98.26	99.16	96.76	97.89	97.94	98.53	93.23	98.65
	Specificity	98.55	99.32	97.39	97.96	97.96	99.12	91.69	99.28
	Sensitivity	98.26	99.16	96.76	97.89	97.94	98.53	93.23	98.65
NR	AUC	97.54	96.44	91.27	92.03	92.85	95.02	90.23	98.38
	Specificity	96.74	97.73	93.77	93.27	93.31	94.02	83.11	98.42
	Sensitivity	97.54	96.44	91.27	92.03	92.85	95.02	90.23	98.38
REM	AUC	95.67	94.48	88.58	91.63	93.83	89.69	87.14	95.27
	Specificity	97.08	98.48	91.56	90.12	93.42	91.54	93.42	93.21
	Sensitivity	95.67	94.48	88.58	91.63	93.83	89.69	87.14	95.27

Table 10.2: Evaluation Metric (%) obtained for mouse data

accuracy obtained for different algorithms and sleep stages in human dataset for CV and IV data arrangement whereas Table 10.4 presents different evaluation metrics for the same dataset in IV data arrangement.

Sleep Stage	DT	RF	kNN			MLP	SVM	XGB
			k = 5	k = 11	k = 19			
	CV							
W	89.29	97.86	98.75	97.86	95.36	97.32	94.11	96.43
N1	88.04	97.41	99.11	98.57	97.77	94.64	92.77	97.41
N2	**93.19**	**99.43**	**99.62**	**99.33**	**99.29**	**99.43**	**97.71**	**99.24**
N3	76.39	82.08	66.94	69.31	70.42	73.61	67.08	78.61
REM	91.79	99.14	98.79	98.21	97.71	93.86	82.29	98.86
	IV							
W	55.78	65.45	67.87	68.03	67.06	71.81	70.64	64.37
N1	57.58	64.72	69.48	69.81	70.29	68.84	70.63	62.86
N2	**70.14**	**75.25**	**75.55**	**75.27**	**75.35**	**80.55**	**80.89**	**73.9**
N3	62.8	64.44	70.91	66.81	69.81	69.35	73.28	63.79
REM	64.91	69.79	72.73	72.82	72.62	77.77	76.85	69.53

Table 10.3: Accuracy (%) obtained for human data

The difference in the metric values across different sleep stages maybe due to the difference in the amount of data present in the different sleep stages. Since the mouse data are recorded

Sleep Stage	Metric	IV							
		DT	RF	kNN			MLP	SVM	XGB
				k = 5	k = 11	k = 19			
W	AUC	55.78	65.45	67.87	68.03	67.06	71.81	70.64	64.37
	Spec	52.23	59.47	56.78	58.35	59.17	62.62	60.53	57.81
	Sen	55.78	65.45	67.87	68.03	67.06	71.81	70.64	64.37
N1	AUC	57.58	64.72	69.48	69.81	70.29	68.84	70.63	62.86
	Spec	53.5	59.85	63.18	64.49	64.76	66.07	66.74	59.05
	Sen	57.58	64.72	69.48	69.81	70.29	68.84	70.63	62.86
N2	AUC	70.14	75.25	75.55	75.27	75.35	80.55	80.89	73.9
	Spec	65.89	71.23	70.62	70.8	71.16	77.71	74.79	71.72
	Sen	70.14	75.25	75.55	75.27	75.35	80.55	80.89	73.9
N3	AUC	62.8	64.44	70.91	66.81	69.18	69.35	73.28	63.79
	Spec	52.67	51.9	71.12	65.52	66.81	67.03	65.09	49.57
	Sen	62.8	64.44	70.91	66.81	69.18	69.35	73.28	63.79
REM	AUC	64.91	69.79	72.73	72.82	72.62	77.77	76.85	69.53
	Spec	59.91	67.74	71.09	71.66	71.37	74.95	73.34	65.74
	Sen	64.91	69.79	72.73	72.82	72.62	77.77	76.85	69.53

Table 10.4: Evaluation Metric (%) obtained for human data

for 24 hours and the human data are PSG recording, the amount of data present in different sleep for each mouse and human subject differs. Along the same line, another variable which is to be considered here is the number of subjects used in different sleep stage. As some sleep stage may not be present in some individual's PSG data, the number of subjects used in training differs. Though this difference is not very high, it cannot be ignored. The reason behind the difference in accuracy obtained in different sleep has to be further investigated. Higher accuracy obtained for mouse dataset can be due to consistent experimental procedure used to induce TBI in mouse as opposed to the heterogeneity of TBI acquired in humans in terms of location of impact and time since impact. The ability of the ML algorithms to learn patterns using the extracted features which enables it to differentiate between TBI and control subjects is showcased by the high accuracy obtained in CV data arrangement. When using a larger data we expect the classification accuracy obtained in IV data arrangement to converge to the results obtained in CV data arrangement thereby, indicating that the model

has learnt more general parameters. Top panel of Fig.10.1 shows the correlation matrix obtained for top features selected using RFE for human N2 stage data used to develop rule based ML models. As it can be observed, most of the selected features are connectivity features such as PLV and coherence and the features are not highly correlated. Bottom panel of Fig.10.1 shows the number of times the respective features were selected in a total on 153 IV iterations.

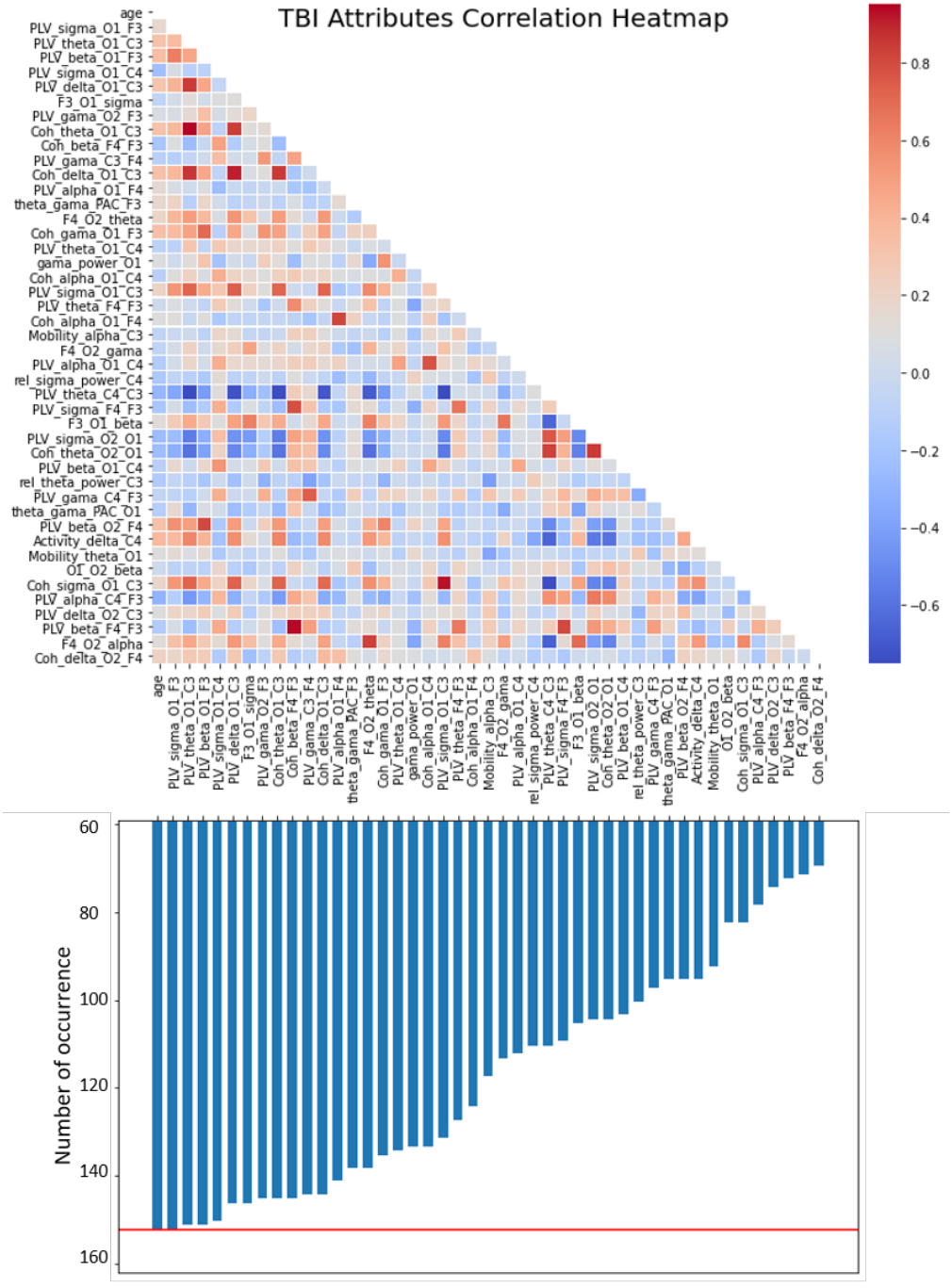

Figure 10.1: Correlation matrix for the most frequently chosen features in N2 sleep stage of human data

Chapter 11

Future work

As it can be observed from the results shown in the previous section, the accuracy obtained for human subjects in IV data arrangement is not optimal. The heterogeneous nature of EEG and the injury which might have led to the above results have been discussed. However, one critical assumption which the traditional ML algorithms make is that the training data and the testing data are drawn from the same distribution. The major hypothesis in traditional ML that the data on which the classifier is trained and the data on the classifier is evaluated belong to the same feature space and follow the same probability distribution often do not hold good in real-world problems. The difference in data distribution typically occurs since the data is acquired from different subjects and one of the ways in which traditional ML overcome this problem is by relying on a massive amount of training data that can account to maximum variability. When the distribution changes, most models need to be retrained from scratch which is very expensive in most cases and may lead to an overfit. One emerging solution to this is transfer learning (TL) [120]. TL aims to learn distributions over different domains and uses the previously learned knowledge on the target task. The potential future work can focus on applying TL/domain adaptation techniques to use the knowledge learnt using the mouse data to accurately classify human EEG.

Bibliography

[1] U. R. Acharya, S. V. Sree, G. Swapna, R. J. Martis, and J. S. Suri. Automated eeg analysis of epilepsy: a review. *Knowledge-Based Systems*, 45:147–165, 2013.

[2] S. Aviyente, E. M. Bernat, W. S. Evans, and S. R. Sponheim. A phase synchrony measure for quantifying dynamic functional integration in the brain. Technical report, Wiley Online Library, 2011.

[3] A. M. Bastos and J.-M. Schoffelen. A tutorial review of functional connectivity analysis methods and their interpretational pitfalls. *Frontiers in systems neuroscience*, 9:175, 2016.

[4] R. Battiti. Using mutual information for selecting features in supervised neural net learning. *IEEE Transactions on neural networks*, 5(4):537–550, 1994.

[5] J. L. Bentley. Multidimensional binary search trees used for associative searching. *Communications of the ACM*, 18(9):509–517, 1975.

[6] N. Bigdely-Shamlo, T. Mullen, C. Kothe, K.-M. Su, and K. A. Robbins. The prep pipeline: standardized preprocessing for large-scale eeg analysis. *Frontiers in neuroinformatics*, 9:16, 2015.

[7] N. Bigdely-Shamlo, J. Touryan, A. Ojeda, C. Kothe, T. Mullen, and K. Robbins. Automated eeg mega-analysis i: Spectral and amplitude characteristics across studies. *NeuroImage*, 207:116361, 2020.

[8] C. M. Bishop. *Pattern recognition and machine learning*. springer, 2006.

[9] B. E. Boser, I. M. Guyon, and V. N. Vapnik. A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational learning theory*, pages 144–152, 1992.

[10] A. P. Bradley. The use of the area under the roc curve in the evaluation of machine learning algorithms. *Pattern recognition*, 30(7):1145–1159, 1997.

[11] L. Breiman. Bagging predictors. *Machine learning*, 24(2):123–140, 1996.

[12] L. Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

[13] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen. *Classification and regression trees*. CRC press, 1984.

[14] D. L. Brody, C. Mac Donald, C. C. Kessens, C. Yuede, M. Parsadanian, M. Spinner, E. Kim, K. E. Schwetye, D. M. Holtzman, and P. V. Bayly. Electromagnetic controlled cortical impact device for precise, graded experimental traumatic brain injury. *Journal of neurotrauma*, 24(4):657–673, 2007.

[15] R. Bruña, F. Maestú, and E. Pereda. Phase locking value revisited: teaching new tricks to an old dog. *Journal of neural engineering*, 15(5):056011, 2018.

[16] P. Bussotti. On the genesis of the lagrange multipliers. *Journal of optimization theory and applications*, 117(3):453–459, 2003.

[17] G. C. Carter. Coherence and time delay estimation. *Proceedings of the IEEE*, 75(2):236–255, 1987.

[18] G. Chandrashekar and F. Sahin. A survey on feature selection methods. *Computers & Electrical Engineering*, 40(1):16–28, 2014.

[19] T. Chen and C. Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pages 785–794, 2016.

[20] X.-P. Chen, L.-Y. Tao, and A. Cn Chen. Electroencephalogram and evoked potential parameters examined in chinese mild head injury patients for forensic medicine. *Neuroscience bulletin*, 22(3):165, 2006.

[21] L. Cohen. *Time-frequency analysis*, volume 778. Prentice Hall PTR Englewood Cliffs, NJ, 1995.

[22] E. Combrisson, T. Nest, A. Brovelli, R. A. A. Ince, J. L. P. Soto, A. Guillot, and K. Jerbi. Tensorpac: An open-source Python toolbox for tensor-based phase-amplitude coupling measurement in electrophysiological brain signals. *PLoS computational biology*, 16(10):e1008302, Oct. 2020.

[23] J. W. Cooley and J. W. Tukey. An algorithm for the machine calculation of complex fourier series. *Mathematics of computation*, 19(90):297–301, 1965.

[24] C. Cortes and V. Vapnik. Support-vector networks. *Machine learning*, 20(3):273–297, 1995.

[25] R. Couronné, P. Probst, and A.-L. Boulesteix. Random forest versus logistic regression: a large-scale benchmark experiment. *BMC bioinformatics*, 19(1):1–14, 2018.

[26] T. Cover and P. Hart. Nearest neighbor pattern classification. *IEEE transactions on information theory*, 13(1):21–27, 1967.

[27] C. DaVBI. Department of defense numbers for traumatic brain injury. 2011.

[28] N. S. Dhillon, A. Sutandi, M. Vishwanath, M. M. Lim, H. Cao, and D. Si. A raspberry pi-based traumatic brain injury detection system for single-channel electroencephalogram. *Sensors*, 21(8):2779, 2021.

[29] R. S. DOW, G. ULETT, and J. RAAF. Electroencephalographic studies immediately following head injury. *American Journal of Psychiatry*, 101(2):174–183, 1944.

[30] M. Faul, L. Xu, M. Wald, and V. Coronado. Centers for disease control and prevention, national center for injury prevention and control, 2010.

[31] T. Fawcett. An introduction to roc analysis. *Pattern recognition letters*, 27(8):861–874, 2006.

[32] R. L. Figueroa, Q. Zeng-Treitler, S. Kandula, and L. H. Ngo. Predicting sample size required for classification performance. *BMC medical informatics and decision making*, 12(1):1–10, 2012.

[33] E. Fix and J. Hodges. An important contribution to nonparametric discriminant analysis and density estimation. *International Statistical Review*, 3(57):233–238, 1951.

[34] C. for Disease Control, Prevention, et al. Report to congress on traumatic brain injury in the united states: epidemiology and rehabilitation. *National Center for Injury Prevention and Control*, pages 1–72, 2015.

[35] N. C. for Injury Prevention and C. (US). *Report to Congress on mild traumatic brain injury in the United States: steps to prevent a serious public health problem.* Centers for Disease Control and Prevention, 2003.

[36] W. J. Freeman. Hilbert transform for brain waves. *Scholarpedia*, 2(1):1338, 2007.

[37] J. Friedman, T. Hastie, R. Tibshirani, et al. *The elements of statistical learning*, volume 1. Springer series in statistics New York, 2001.

[38] L. J. Gabard-Durnam, A. S. Mendez Leal, C. L. Wilkinson, and A. R. Levin. The harvard automated processing pipeline for electroencephalography (happe): standardized processing software for developmental and high-artifact data. *Frontiers in neuroscience*, 12:97, 2018.

[39] T. Gasser, P. Bächer, and J. Möcks. Transformations towards the normal distribution of broad band spectral parameters of the eeg. *Electroencephalography and clinical neurophysiology*, 53(1):119–124, 1982.

[40] N. Gosselin, M. Lassonde, D. Petit, S. Leclerc, V. Mongrain, A. Collie, and J. Montplaisir. Sleep following sport-related concussions. *Sleep medicine*, 10(1):35–46, 2009.

[41] A. Gramfort, M. Luessi, E. Larson, D. A. Engemann, D. Strohmeier, C. Brodbeck, R. Goj, M. Jas, T. Brooks, L. Parkkonen, et al. Meg and eeg data analysis with mne-python. *Frontiers in neuroscience*, 7:267, 2013.

[42] A. Gramfort, M. Luessi, E. Larson, D. A. Engemann, D. Strohmeier, C. Brodbeck, R. Goj, M. Jas, T. Brooks, L. Parkkonen, and M. S. Hämäläinen. MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*, 7(267):1–13, 2013.

[43] I. Guyon and A. Elisseeff. An introduction to variable and feature selection. *Journal of machine learning research*, 3(Mar):1157–1182, 2003.

[44] M. A. Hall. Correlation-based feature selection of discrete and numeric class machine learning. 2000.

[45] B. J. He, J. M. Zempel, A. Z. Snyder, and M. E. Raichle. The temporal structures and functional significance of scale-free brain activity. *Neuron*, 66(3):353–369, 2010.

[46] J. Hérault and B. Ans. Neuronal network with modifiable synapses: decoding of composite sensory messages under unsupervised and permanent learning. *Comptes rendus de l'Academie des sciences. Serie III, Sciences de la vie*, 299(13):525–528, 1984.

[47] J. Hérault, C. Jutten, and B. Ans. Détection de grandeurs primitives dans un message composite par une architecture de calcul neuromimétique en apprentissage non supervisé. In *10 Colloque sur le traitement du signal et des images, FRA, 1985*. GRETSI, Groupe d'Etudes du Traitement du Signal et des Images, 1985.

[48] B. Hjorth. Eeg analysis based on time domain properties. *Electroencephalography and clinical neurophysiology*, 29(3):306–310, 1970.

[49] T. K. Ho. Random decision forests. In *Proceedings of 3rd international conference on document analysis and recognition*, volume 1, pages 278–282. IEEE, 1995.

[50] T. K. Ho. The random subspace method for constructing decision forests. *IEEE transactions on pattern analysis and machine intelligence*, 20(8):832–844, 1998.

[51] T. K. Ho. A data complexity analysis of comparative advantages of decision forest constructors. *Pattern Analysis & Applications*, 5(2):102–112, 2002.

[52] A. E. Hoerl and R. W. Kennard. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1):55–67, 1970.

[53] J. Hua, Z. Xiong, J. Lowey, E. Suh, and E. R. Dougherty. Optimal number of features as a function of sample size for various classification rules. *Bioinformatics*, 21(8):1509–1515, 2005.

[54] J. D. Hunter. Matplotlib: A 2d graphics environment. *Computing in Science & Engineering*, 9(3):90–95, 2007.

[55] A. Hyvarinen. Fast and robust fixed-point algorithms for independent component analysis. *IEEE transactions on Neural Networks*, 10(3):626–634, 1999.

[56] A. Hyvärinen and E. Oja. Independent component analysis: algorithms and applications. *Neural networks*, 13(4-5):411–430, 2000.

[57] A. Hyvtirinen, J. Karhunen, and E. Oja. Independent component analysis. 2001.

[58] J. N. Ianof and R. Anghinah. Traumatic brain injury: An eeg point of view. *Dementia & neuropsychologia*, 11(1):3–5, 2017.

[59] T. Inouye, K. Shinosaki, H. Sakamoto, S. Toi, S. Ukai, A. Iyama, Y. Katsuda, and M. Hirano. Quantification of eeg irregularity by use of the entropy of the power spectrum. *Electroencephalography and clinical neurophysiology*, 79(3):204–210, 1991.

[60] A. Jacquin, S. Kanakia, D. Oberly, and L. S. Prichep. A multimodal biomarker for concussion identification, prognosis and management. *Computers in biology and medicine*, 102:95–103, 2018.

[61] A. K. Jain and W. G. Waller. On the optimal number of features in the classification of multivariate gaussian data. *Pattern recognition*, 10(5-6):365–374, 1978.

[62] G. James, D. Witten, T. Hastie, and R. Tibshirani. *An introduction to statistical learning*, volume 112. Springer, 2013.

[63] J. Jeong. Eeg dynamics in patients with alzheimer's disease. *Clinical neurophysiology*, 115(7):1490–1505, 2004.

[64] E. John. Normative data bank and neurometrics. basic concepts, methods and results of norm constructions. *Methods od analysis of brain electrical and magnetic signals. EEG handbook*, 1:449–498, 1987.

[65] E. John, H. Ahn, L. Prichep, M. Trepetin, D. Brown, and H. Kaye. Developmental equations for the electroencephalogram. *Science*, 210(4475):1255–1258, 1980.

[66] E. R. John, L. Prichep, J. Fridman, and P. Easton. Neurometrics: computer-assisted differential diagnosis of brain dysfunctions. *Science*, 239(4836):162–169, 1988.

[67] M. Junghöfer, T. Elbert, D. M. Tucker, and B. Rockstroh. Statistical control of artifacts in dense array eeg/meg studies. *Psychophysiology*, 37(4):523–532, 2000.

[68] E. S. Kenzie, E. L. Parks, E. D. Bigler, M. M. Lim, J. C. Chesnutt, and W. Wakeland. Concussion as a multi-scale complex system: an interdisciplinary synthesis of current knowledge. *Frontiers in neurology*, 8:513, 2017.

[69] R. Kohavi et al. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Ijcai*, volume 14, pages 1137–1145. Montreal, Canada, 1995.

[70] R. Kohavi and G. H. John. Wrappers for feature subset selection. *Artificial intelligence*, 97(1-2):273–324, 1997.

[71] C. A. E. Kothe and T.-p. Jung. Artifact removal techniques with signal reconstruction, Apr. 28 2016. US Patent App. 14/895,440.

[72] S. Kullback and R. A. Leibler. On information and sufficiency. *The annals of mathematical statistics*, 22(1):79–86, 1951.

[73] J.-P. Lachaux, E. Rodriguez, J. Martinerie, and F. J. Varela. Measuring phase synchrony in brain signals. *Human brain mapping*, 8(4):194–208, 1999.

[74] J. L. Lagrange. *Mécanique analytique*, volume 1. Mallet-Bachelier, 1853.

[75] W.-Y. Loh. Classification and regression trees. *Wiley interdisciplinary reviews: data mining and knowledge discovery*, 1(1):14–23, 2011.

[76] W.-Y. Loh. Fifty years of classification and regression trees. *International Statistical Review*, 82(3):329–348, 2014.

[77] M. Matoušek and I. Petersén. Automatic evaluation of eeg background activity by means of age-dependent eeg quotients. *Electroencephalography and clinical neurophysiology*, 35(6):603–612, 1973.

[78] M. H. Modarres, N. N. Kuzma, T. Kretzmer, A. I. Pack, and M. M. Lim. Eeg slow waves in traumatic brain injury: Convergent findings in mouse and man. *Neurobiology of sleep and circadian rhythms*, 2:59–70, 2017.

[79] M. H. Modarres, R. A. Opel, K. B. Weymann, and M. M. Lim. Strong correlation of novel sleep electroencephalography coherence markers with diagnosis and severity of posttraumatic stress disorder. *Scientific reports*, 9(1):1–10, 2019.

[80] T. T. Munia and S. Aviyente. Time-frequency based phase-amplitude coupling measure for neuronal oscillations. *Scientific reports*, 9(1):1–15, 2019.

[81] T. T. Munia, A. Haider, C. Schneider, M. Romanick, and R. Fazel-Rezai. A novel eeg based spectral analysis of persistent brain function alteration in athletes with concussion history. *Scientific reports*, 7(1):1–13, 2017.

[82] T. Nakamura, F. G. Hillary, and B. B. Biswal. Resting network plasticity following brain injury. *PloS one*, 4(12):e8220, 2009.

[83] H. Nolan, R. Whelan, and R. B. Reilly. Faster: fully automated statistical thresholding for eeg artifact rejection. *Journal of neuroscience methods*, 192(1):152–162, 2010.

[84] P. L. Nunez, R. Srinivasan, et al. *Electric fields of the brain: the neurophysics of EEG*. Oxford University Press, USA, 2006.

[85] P. L. Nunez, R. Srinivasan, A. F. Westdorp, R. S. Wijesinghe, D. M. Tucker, R. B. Silberstein, and P. J. Cadusch. Eeg coherency: I: statistics, reference electrode, volume conduction, laplacians, cortical imaging, and interpretation at multiple scales. *Electroencephalography and clinical neurophysiology*, 103(5):499–515, 1997.

[86] M. R. Nuwer, D. A. Hovda, L. M. Schrader, and P. M. Vespa. Routine and quantitative eeg in mild traumatic brain injury. *Clinical Neurophysiology*, 116(9):2001–2025, 2005.

[87] S. M. Omohundro. *Five balltree construction algorithms*. International Computer Science Institute Berkeley, 1989.

[88] A. V. Oppenheim. *Discrete-time signal processing*. Pearson Education India, 1999.

[89] H. J. Orff, L. Ayalon, and S. P. Drummond. Traumatic brain injury and sleep disturbance: a review of current research. *The Journal of head trauma rehabilitation*, 24(3):155–165, 2009.

[90] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[91] J. Platt et al. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. *Advances in large margin classifiers*, 10(3):61–74, 1999.

[92] D. M. Powers. Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation. *arXiv preprint arXiv:2010.16061*, 2020.

[93] L. Prichep, E. John, S. Ferris, L. Rausch, Z. Fang, R. Cancro, C. Torossian, and B. Reisberg. Prediction of longitudinal cognitive decline in normal elderly with subjective complaints using electrophysiological imaging. *Neurobiology of aging*, 27(3):471–481, 2006.

[94] L. S. Prichep, A. Jacquin, J. Filipenko, S. G. Dastidar, S. Zabele, A. Vodencarevic, and N. S. Rothman. Classification of traumatic brain injury severity using informed data reduction in a series of binary classifier algorithms. *IEEE transactions on neural systems and rehabilitation engineering*, 20(6):806–822, 2012.

[95] J. R. Quinlan. Induction of decision trees. *Machine learning*, 1(1):81–106, 1986.

[96] J. R. Quinlan. *C4. 5: programs for machine learning*. Elsevier, 2014.

[97] A. Ramakrishnan and J. Satyanarayana. Reconstruction of eeg from limited channel acquisition using estimated signal correlation. *Biomedical Signal Processing and Control*, 27:164–173, 2016.

[98] P. E. Rapp, D. O. Keyser, A. Albano, R. Hernandez, D. B. Gibson, R. A. Zambon, W. D. Hairston, J. D. Hughes, A. Krystal, and A. S. Nichols. Traumatic brain injury detection using electrophysiological methods. *Frontiers in human neuroscience*, 9:11, 2015.

[99] J. J. Renger, S. L. Dunn, S. L. Motzel, C. Johnson, and K. S. Koblan. Sub-chronic administration of zolpidem affects modifications to rat sleep architecture. *Brain research*, 1010(1-2):45–54, 2004.

[100] K. A. Robbins, J. Touryan, T. Mullen, C. Kothe, and N. Bigdely-Shamlo. How sensitive are eeg results to preprocessing methods: a benchmarking study. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 28(5):1081–1090, 2020.

[101] Y. Roy, H. Banville, I. Albuquerque, A. Gramfort, T. H. Falk, and J. Faubert. Deep learning-based electroencephalography analysis: a systematic review. *Journal of neural engineering*, 16(5):051001, 2019.

[102] A. L. Samuel. Some studies in machine learning using the game of checkers. *IBM Journal of research and development*, 3(3):210–229, 1959.

[103] D. K. Sandsmark, J. E. Elliott, and M. M. Lim. Sleep-wake disturbances after traumatic brain injury: synthesis of human and animal studies. *Sleep*, 40(5), 2017.

[104] H. Schütze, C. D. Manning, and P. Raghavan. *Introduction to information retrieval*, volume 39. Cambridge University Press Cambridge, 2008.

[105] C. E. Shannon. A mathematical theory of communication. *ACM SIGMOBILE mobile computing and communications review*, 5(1):3–55, 2001.

[106] M. Sokolova, N. Japkowicz, and S. Szpakowicz. Beyond accuracy, f-score and roc: a family of discriminant measures for performance evaluation. In *Australasian joint conference on artificial intelligence*, pages 1015–1021. Springer, 2006.

[107] M. Sokolova and G. Lapalme. A systematic analysis of performance measures for classification tasks. *Information processing & management*, 45(4):427–437, 2009.

[108] S. R. Sponheim, K. A. McGuire, S. S. Kang, N. D. Davenport, S. Aviyente, E. M. Bernat, and K. O. Lim. Evidence of disrupted functional connectivity in the brain after combat-related blast injury. *Neuroimage*, 54:S21–S29, 2011.

[109] C. J. Stam, B. Jones, G. Nolte, M. Breakspear, and P. Scheltens. Small-world networks and functional connectivity in alzheimer's disease. *Cerebral cortex*, 17(1):92–99, 2007.

[110] G. Teasdale and B. Jennett. Assessment of coma and impaired consciousness: a practical scale. *The Lancet*, 304(7872):81–84, 1974.

[111] N. V. Thakor and S. Tong. Advances in quantitative electroencephalogram analysis methods. *Annu. Rev. Biomed. Eng.*, 6:453–495, 2004.

[112] A. Tharwat. Classification assessment methods. *Applied Computing and Informatics*, 2020.

[113] R. Thatcher, C. Biver, J. Gomez, D. North, R. Curtin, R. Walker, and A. Salazar. Estimation of the eeg power spectrum using mri t2 relaxation time in traumatic brain injury. *Clinical Neurophysiology*, 112(9):1729–1745, 2001.

[114] R. Thatcher, D. Cantor, R. McAlaster, F. Geisler, and P. Krause. Comprehensive predictions of outcome in closed head-injured patients: The development of prognostic equations. *Annals of the New York Academy of Sciences*, 620(1):82–101, 1991.

[115] R. Thatcher, R. Walker, and S. Giudice. Human cerebral hemispheres develop at different rates and ages. *Science*, 236(4805):1110–1113, 1987.

[116] R. W. Thatcher. Normative eeg databases and eeg biofeedback. *Journal of Neurotherapy*, 2(4):8–39, 1998.

[117] R. W. Thatcher and J. F. Lubar. History of the scientific standards of qeeg normative databases. *Introduction to quantitative EEG and neurofeedback: Advanced theory and applications*, 2:29–59, 2009.

[118] R. W. Thatcher, R. Walker, I. Gerson, and F. Geisler. Eeg discriminant analyses of mild head trauma. *Electroencephalography and clinical neurophysiology*, 73(2):94–106, 1989.

[119] K. Thornton. The electrophysiological effects of a brain injury on auditory memory functioning the qeeg correlates of impaired memory. *Archives of Clinical Neuropsychology*, 18(4):363–378, 2003.

[120] S. Thrun and L. Pratt. Learning to learn: Introduction and overview. In *Learning to learn*, pages 3–17. Springer, 1998.

[121] R. Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288, 1996.

[122] A. B. Tort, R. Komorowski, H. Eichenbaum, and N. Kopell. Measuring phase-amplitude coupling between neuronal oscillations of different frequencies. *Journal of neurophysiology*, 104(2):1195–1210, 2010.

[123] R. Trevethan. Sensitivity, specificity, and predictive values: foundations, pliabilities, and pitfalls in research and practice. *Frontiers in public health*, 5:307, 2017.

[124] V. N. Vapnik. *Statistical Learning Theory*. Wiley-Interscience, 1998.

[125] M. Vishwanath, S. Jafarlou, I. Shin, N. Dutt, A. M. Rahmani, M. M. Lim, and H. Cao. Classification of electroencephalogram in a mouse model of traumatic brain injury using machine learning approaches. In *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pages 3335–3338. IEEE, 2020.

[126] M. Vishwanath, S. Jafarlou, I. Shin, M. M. Lim, N. Dutt, A. M. Rahmani, and H. Cao. Investigation of machine learning approaches for traumatic brain injury classification via eeg assessment in mice. *Sensors*, 20(7):2027, 2020.

[127] N. Vivaldi, M. Caiola, K. Solarana, and M. Ye. Evaluating performance of eeg data-driven machine learning for traumatic brain injury classification. *IEEE Transactions on Biomedical Engineering*, 2021.

[128] B. E. Wallace, A. K. Wagner, E. P. Wagner, and J. T. McDeavitt. A history and review of quantitative electroencephalography in traumatic brain injury. *The Journal of head trauma rehabilitation*, 16(2):165–190, 2001.

[129] M. L. Waskom. seaborn: statistical data visualization. *Journal of Open Source Software*, 6(60):3021, 2021.

[130] D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world'networks. *nature*, 393(6684):440–442, 1998.

[131] P. Welch. The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. *IEEE Transactions on audio and electroacoustics*, 15(2):70–73, 1967.

[132] A. Widmann, E. Schröger, and B. Maess. Digital filter design for electrophysiological data–a practical approach. *Journal of neuroscience methods*, 250:34–46, 2015.

[133] B. R. Williams, S. E. Lazic, and R. D. Ogilvie. Polysomnographic and quantitative eeg analysis of subjects with long-term insomnia complaints associated with mild traumatic brain injury. *Clinical Neurophysiology*, 119(2):429–438, 2008.

[134] J. T. Willie, M. M. Lim, R. E. Bennett, A. A. Azarion, K. E. Schwetye, and D. L. Brody. Controlled cortical impact traumatic brain injury acutely disrupts wakefulness and extracellular orexin dynamics as determined by intracerebral microdialysis in mice. *Journal of neurotrauma*, 29(10):1908–1921, 2012.

[135] I. Winkler, S. Haufe, and M. Tangermann. Automatic classification of artifactual ica-components for artifact removal in eeg signals. *Behavioral and Brain Functions*, 7(1):1–15, 2011.

[136] D. Yael, J. J. Vecht, and I. Bar-Gad. Filter-based phase shifts distort neuronal timing information. *Eneuro*, 5(2), 2018.