

Estimation of cognitive state improvement using EEG brain  
imaging after tDCS stimulation therapy



Author

MUDASSAR AYUB

Regn # 00000204075

Supervisor

DR. M. JAWAD KHAN

DEPARTMENT OF ROBOTICS AND ARTIFICIAL INTELLIGENCE  
SCHOOL OF MECHANICAL & MANUFACTURING ENGINEERING  
NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY  
ISLAMABAD

June 2021

Estimation of cognitive state improvement using EEG brain imaging  
after tDCS stimulation therapy

Author

MUDASSAR AYUB

Regn # 00000204075

A thesis submitted in partial fulfillment of the requirements for the degree of  
MS Robotics and Intelligent Machines Engineering

Thesis Supervisor:

Dr. M. JAWAD KHAN

Thesis Supervisor's Signature:

---

DEPARTMENT OF ROBOTICS AND ARTIFICIAL INTELLIGENCE  
SCHOOL OF MECHANICAL & MANUFACTURING ENGINEERING  
NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY,  
ISLAMABAD

June 2021

National University of Sciences and Technology

**MASTER THESIS WORK**

We hereby recommend that the dissertation prepared under our supervision by: **Mr. Mudassar Ayub Regn#00000204075** Titled: **“Estimation of cognitive state improvement Using EEG brain imaging after tDCS stimulation therapy”** be accepted in partial fulfillment of the requirements for the award of **MS Robotics & Intelligent Machine Engineering** degree. (Grade\_\_\_\_)

**Examination Committee Members**

1. Name: Dr. Yasar Ayaz Signature: \_\_\_\_\_
2. Name: Dr. Asim Waris Signature: \_\_\_\_\_
3. Name: Dr. Syed Omer Gillani Signature: \_\_\_\_\_

Supervisor's name: Dr. M. Jawad Khan Signature: \_\_\_\_\_

Date: \_\_\_\_\_

\_\_\_\_\_

Head of Department

\_\_\_\_\_

Date

**COUNTERSIGNED**

Date: \_\_\_\_\_

\_\_\_\_\_

Principal

## Thesis Acceptance Certificate

It is certified that the final copy of MS Thesis written by *Mudassar Ayub* (Registration No. 00000204075), of Department of Robotics and Intelligent Machine Engineering (SMME) has been vetted by undersigned, found complete in all respects as per NUST statutes / regulations, is free from plagiarism, errors and mistakes and is accepted as a partial fulfilment for award of MS Degree. It is further certified that necessary amendments as pointed out by GEC members of the scholar have also been incorporated in this dissertation.

Signature: \_\_\_\_\_

Name of Supervisor: Dr. M. Jawad Khan (Supervisor)

Date: \_\_\_\_\_

Signature (HOD): \_\_\_\_\_

Date: \_\_\_\_\_

Signature (Principal): \_\_\_\_\_

Date: \_\_\_\_\_

## **Declaration**

I certify that this research work titled “*Estimation of cognitive state improvement Using EEG brain imaging after tDCS stimulation therapy*” is my own work. The work has not been presented elsewhere for assessment. The material that has been used from other sources it has been properly acknowledged / referred.

MUDASSAR AYUB

Reg# 00000204075

## **Plagiarism Certificate (Turnitin Report)**

This thesis has been checked for Plagiarism. Turnitin report endorsed by Supervisor is attached.

MUDASSAR AYUB

Regn # 00000204075

DR. M. JAWAD KHAN

(Supervisor)

## **Copyright Statement**

- Copyright in text of this thesis rests with the student author. Copies (by any process) either in full, or of extracts, may be made only in accordance with instructions given by the author and lodged in the Library of NUST School of Mechanical & Manufacturing Engineering (SMME). Details may be obtained by the Librarian. This page must form part of any such copies made. Further copies (by any process) may not be made without the permission (in writing) of the author.
- The ownership of any intellectual property rights which may be described in this thesis is vested in NUST School of Mechanical & Manufacturing Engineering, subject to any prior agreement to the contrary, and may not be made available for use by third parties without the written permission of the SMME, which will prescribe the terms and conditions of any such agreement.
- Further information on the conditions under which disclosures and exploitation may take place is available from the Library of NUST School of Mechanical & Manufacturing Engineering, Islamabad.

## **Acknowledgements**

I am thankful to Allah Almighty for bestowing me with the energy and will that enabled me to successfully complete this project.

I am thankful to my supervisor Dr. M. Jawad Khan and the School of Mechanical and Manufacturing Engineering who supported me in every possible way throughout the course of my thesis.

*Dedicated to my parents and friends without whom support I would  
not have achieved such an accomplishment.*

## Abstract

Brain is the most important organ in human body. The effects of brain impairment are wide ranging and includes cognition, fatigue, sleep issues, headaches, dizziness, impaired self-awareness, clinical depression, attention and concentration problems, epilepsy, struggling in making decisions and many more. Brain rehabilitation via noninvasive techniques such as tDCS helps patients relearn functions and improvement in cognition, lost as a result of a brain impairment. Transcranial direct current stimulation (tDCS) is one such noninvasive, safe and convenient neuro-modulatory technique in neurological rehabilitation, treatment, and other aspects of brain disorders. The efficacy of tDCS in estimation of brain cognitive state improvement using changes in small electrical brain voltages recorded by Electroencephalogram (EEG) of 10 subjects is assessed by applying Event related Desynchronization (ERDs) to Motor Imagery Period (MIP) and Rest Period (RP) of pre stimulation and post stimulation data and features extraction technique such as common spatial pattern (CSP). The results suggest a decrease of contralateral ERDs oscillatory activity related to an event as per the hypothesis in anodal post stimulation than pre stimulation across all channels for six subjects. Further Linear Discrimination Analysis (LDA) a machine learning model applied on CSP and ERDs proved that the classification accuracy between Motor imagery period (MIP) and Rest period (RP) after the stimulation therapy is higher than the Pre stimulation Motor Imagery period (MIP) and Rest period (RP).

**Keywords:** *tDCS Transcranial Direct Current Stimulations, EEG Electroencephalogram, ERD Event related Desynchronization, CSP Common Spatial Pattern, MIP Motor Imagery Period, RP Rest Period, LDA Linear Discriminant Analysis*

## Table of Contents

FORM TH-4.....	I
Thesis Acceptance Certificate.....	II
Declaration.....	III
Plagiarism Certificate (Turnitin Report).....	IV
Copyright Statement.....	V
Acknowledgements.....	VI
Abstract.....	VIII
Table of Contents.....	IX
Table of Figures.....	X
List of Tables.....	XII
List of Abbreviations.....	XIII
Chapter 1: INTRODUCTION.....	1
1.1 Aims and Objectives:.....	1
1.2 Research Methodology:.....	2
Chapter 2: Literature Review.....	3
2.1 Understanding the EEG and tDCS.....	3
2.1.1 Electroencephalograph (EEG).....	3
2.1.1.1 Motion Artifacts.....	4
2.1.1.2 Instrumental Artifacts.....	4
2.1.1.3 Physiological Artifacts.....	4
2.1.1.4 Eye Blink Artifacts.....	4
2.1.2 Transcranial Direct Current Stimulation (tDCS).....	5
2.2 EEG and tDCS for Brain Rehabilitation.....	6
Chapter 3: METHODOLOGY.....	8
3.1 Proposed Scheme.....	8
3.2 Data Collection.....	8
3.2.1 Dataset 1 Experimental Paradigm.....	9
3.2.2 Dataset 2 Experimental Paradigm.....	14
3.3 Data Preprocessing:.....	16
3.3.1 Data Labeling.....	16
3.3.2 Intra Correlation.....	17
3.3.3 Data Standardization.....	19
3.4 Data Filter.....	19
3.4.1 ICA.....	20
3.4.2 ICA components Rejection.....	24
3.5 Event Related Potential (ERP)/Event Related Desynchronization (ERD).....	25

3.6 Common Spatial Patterns (CSP) .....	27
3.7 Linear Discriminant Analysis (LDA) .....	28
Chapter 4: RESULTS AND DISCUSSIONS .....	29
4.1 Results of Data labeling accuracy of Pre stimulation and Post stimulation data .....	29
4.2 Common Spatial Pattern (CSP).....	33
4.2.1 Subject 10 CSP.....	35
4.3 Event Related De synchronization (ERD) .....	36
4.3.1 ERDs Channels Cz, C2, C4, C6.....	37
4.3.1.1 Subject 2.....	37
4.3.1.2 Subject 9.....	37
4.3.1.3 Subject 4.....	38
4.3.1.4 Subject 7.....	39
4.3.1.5 Subject 10.....	40
4.3.2 Channels FCz, FC2, FC4, FC6.....	42
4.3.2.1 Subject 1 ERDs .....	42
4.3.2.2 Subject 3.....	42
4.3.2.3 Subject 9.....	43
4.3.2.4 Subject 7.....	44
4.3.2.5 Subject 10.....	45
4.4 Linear Discriminant Analysis (LDA) Classification .....	46
Chapter 5: CONCLUSIN AND FUTURE WORK .....	53
5.1 Conclusion .....	53
5.2 Future Work.....	53
APPENDIX I ERD.....	54
APPENDIX II CSP.....	59
REFERENCES .....	71

### **Table of Figures**

Figure 3-1: Proposed Scheme .....	8
Figure 3-2: Recording Setup.....	11
Figure 3-3: Subject Recording during Experimental Setup.....	12
Figure 3-4:Experimental Setup .....	15
Figure 3-5: ICA components with heat maps, segmented ERP and spectrum .....	24
Figure 3-6: The signal before and after cleaning the data through ICA .....	25
Figure 3-7: EEG signal with stimulus onset .....	26
Figure 3-8: LDA classification .....	28
Figure 4-1: Scatter Plots of Cubic SVM model.....	31
Figure 4-2: Confusion Matrix of Cubic SVM model.....	32
Figure 4-3: True Positive and Negative Rate of Cubic SVM model.....	32
Figure 4-4: Positive Predictive Values and False discovery rates of Cubic SVM model.....	33

Figure 4-5: CSP of pre stimulation and post stimulation.....	34
Figure 4-6: CSP of pre stimulation and post stimulation.....	34
Figure 4-7: CSP of pre stimulation and post stimulation.....	35
Figure 4-8: CSP of pre stimulation and post stimulation.....	35
Figure 4-9: CSP of pre stimulation and post stimulation.....	36
Figure 4-10: CSP of pre stimulation and post stimulation.....	36
Figure 4-11: Pre stimulation ERD for subject 2 .....	37
Figure 4-12: Post stimulation ERD for subject 2.....	37
Figure 4-13: Pre stimulation ERD for subject 9 .....	38
Figure 4-14: Post stimulation ERD for subject 9.....	38
Figure 4-15: Pre stimulation ERD for subject 4 .....	39
Figure 4-16: Post stimulation ERD for subject 4.....	39
Figure 4-17: Pre stimulation ERD for subject 7 .....	40
Figure 4-18: Post stimulation ERD for subject 7.....	40
Figure 4-19: Pre stimulation ERD for subject 10 .....	41
Figure 4-20: Post stimulation ERD for subject 10.....	41
Figure 4-21: Pre stimulation ERD for subject 1 .....	42
Figure 4-22: Post stimulation ERD for subject 1.....	42
Figure 4-23: Pre stimulation ERD for subject 3. ....	43
Figure 4-24: Post stimulation ERD for subject 3.....	43
Figure 4-25: Pre stimulation ERD for subject 9 .....	44
Figure 4-26: Post stimulation ERD for subject 9.....	44
Figure 4-27: Pre stimulation ERD for subject 7 .....	45
Figure 4-28: Post stimulation ERD for subject 7.....	45
Figure 4-29: Pre stimulation ERD for subject 10 .....	46
Figure 4-30: Post stimulation ERD for subject 10.....	46
Figure 4-31: Subject 2 MI Classification Accuracy .....	48
Figure 4-32: Subject 1 MI Classification Accuracy .....	48
Figure 4-33: Subject 9 MI Classification Accuracy .....	49
Figure 4-34: Subject 4 MI Classification Accuracy .....	49
Figure 4-35: MI classification comparison on CSP.....	50
Figure 4-36: Subject 1 MI Classification Accuracy .....	51
Figure 4-37: Subject 9 MI Classification Accuracy .....	51
Figure 4-38: MI classification comparison on CSP.....	52

## List of Tables

Table 3-1: Participants of the study .....	9
Table 3-2: Recording Montage. ....	10
Table 3-3: Pre-Stimulation Motor Imagery Period (MIP) & Rest Period (RP).....	12
Table 3-4: During Stimulation Motor Imagery Period (MIP) & Rest Period (RP) .....	13
Table 3-5: Post Stimulation Motor Imagery Period (MIP) & Rest Period (RP).....	13
Table 3-6: Pre stimulation time MI and RP Intra Correlation Matrix .....	17
Table 3-7: Post stimulation time MI and RP Intra Correlation Matrix.....	18
Table 4-1: Summary of input data for Simple label Classification .....	29
Table 4-2: Summary of different models used for classification.....	29
Table 4-3: Classification Accuracy of Cubic and Fine KNN models.....	31
Table 4-4: LDA Motor Imagery Detection Accuracy dataset 2 .....	47
Table 4-5: LDA Motor Imagery Detection Accuracy dataset 1 .....	50

## List of Abbreviations

<b>Abbreviation</b>	<b>Description</b>
<b>tDCS</b>	Transcranial Direct Current Stimulation
<b>EEG</b>	Electroencephalogram
<b>MIP</b>	Motor Imagery Period
<b>RP</b>	Rest Period
<b>ERD</b>	Event Related De synchronization
<b>CSP</b>	Common Spatial Pattern
<b>LDA</b>	Linear Discriminant Analysis
<b>ICA</b>	Independent Component Analysis

## CHAPTER 1: INTRODUCTION

According to World Health Organization (WHO) Major mental disorders in Pakistan are depression (6%), schizophrenia (1.5%) and epilepsy (1-2%). Depression affects more than 350 million of ages, in all communities [1]. It has significantly contributed to the global burden of disease and is one of the major contributor to a high suicide rate. Mental health problems in developing countries like Pakistan, have in the last few decades reached an appalling level (David DB,2000 Gadit AAM,2005) and are linked to both the current violence in society (Khalily TM ,2011 Khalily MT., 2010) and disruption in its social structure (Gadit AAM,1999).The health care treatment system's response to these problems is Worse in developing countries, the number of psychiatrists and psychiatric beds per head of population is much smaller and the treatment is expensive Fundamentally, there is no established model for mental health care in most developing countries and the majority of psychiatric patients thus seek treatment from non-professional healers using psycho [1].

Brain is the most important organ in human body. The physical effects of brain injury are wide ranging and includes cognition, fatigue, sleep issues, headaches, dizziness, domestic violence, Impaired self-awareness, self-centeredness, impulsive behavior, anger, clinical Depression, personality, behavior, attention and concentration problems, sensory and perceptual problems, epilepsy, motivation and initiation (adynamia), difficulty with making decisions, perseveration (repetition), panic attacks and hearing problems. Brain rehabilitation helps patients relearn functions lost as a result of a brain injury. Transcranial direct current stimulation (tDCS) is a noninvasive, safe and convenient neuro-modulatory technique in neurological rehabilitation, treatment, and other aspects of brain disorders [2]. We aimed to evaluate the effects of tDCS on estimation of brain cognitive state improvement using changes in small electrical brain voltages recorded by Electroencephalogram (EEG) and applying Event related De synchronization data analysis technique and features extraction technique such as common spatial pattern (CSP) over the recorded brain imaging data.

### 1.1 Aims and Objectives:

The aims and objectives of this work are

- Design of Therapy
- Collecting pre and post tDCS simulation EEG data.

- Data preprocessing and artifact removal.
- Estimate cognitive improvement from Data.

## **1.2 Research Methodology:**

The study will be split into the following major objectives to achieve the goal of this research:

### **1.2.1 Literature Review**

This review will discuss the combination of already existing different noninvasive brain stimulation therapy approaches and noninvasive brain recording techniques used for the rehabilitation of brain disorder patients for this research. Testing methods and other requirements will also be researched and discussed in this review.

### **1.2.2 Dataset and Methodology:**

In methodology, there will be a discussion about the two data set used in this research one which is the recorded data set in the lab as a part of research and the other online available dataset. Data preprocessing of bot data set and preprocessing results are also explained in this section

### **1.2.3 • Results:**

In the result section, results of the different classification techniques used in this research is presented.

### **1.2.4 • Conclusion:**

In the end, there will be conclusions regarding this research and future works.

## **CHAPTER 2: LITERATURE REVIEW**

### **2.1 Understanding the EEG and tDCS**

In this study efficacy of tDCS therapy is analyzed on the brain signal acquired through EEG brain imaging. Therefore, it is essential to understand EEG signals and the stimulation therapy.

#### **2.1.1 Electroencephalograph (EEG)**

There are different ways in which for capturing the brain's activity [3]: electrooculography (EOG), electroencephalography (EEG), near-infrared spectroscopy (NIRS), and function near-infrared spectroscopy (fNIRS) are the most recent one.

Electroencephalography is the neurophysiological measurement of electrical activity in the brain as recorded by electrodes placed on the scalp or, in special cases, subdural or in the cerebral cortex. The resulting traces are known as an electroencephalogram (EEG) and represent a summation of post-synaptic potentials from a large number of neurons. [4]

The use of EEG in neuroscience research delivers a number of benefits. EEG is non-invasive in nature, ease of use, portable and cheap for the research subject. Furthermore, the need to restrict the subject's movements is clearly lower than in other fields of neuroscience such as functional magnetic resonance imaging (fMRI). A further benefit is that many EEG applications record spontaneous brain activity, which means that the subject does not need to be able to cooperate with the researcher (as is necessary, for instance, during behavioral testing in neuropsychology) [4]. EEGs have a high temporal resolution compared with techniques such as fMRI and PET and are capable of detecting changes in electrical activity in the brain on a time scale in the millisecond region which means the detection of hemodynamic brain response, spatial resolution is higher than the detection of neural stimulus-response in EEG [5], [6]

In conventional scalp EEG, the recording is obtained by applying electrodes to the scalp using a conductive gel or paste, usually after preparing the scalp area by light abrasion to reduce electrode-scalp impedance. Many systems typically use electrodes which are each attached to an individual wire. Some systems use caps in which electrodes are embedded. This latter method is particularly common when high-density arrays of electrodes are required.

In addition to internal artifacts such as those produced by blinking, there are many artifacts which originate from outside the patient. Following are the types of artifacts present in an EEG signal.

#### **2.1.1.1 Motion Artifacts**

Movement by the patient generates huge artifacts. Different motion artifacts occur in the form of peaks and changes in baseline in data the major reason is due to muscle movements that can occur during the recording of signals and due to improper scalp attachment, they should be removed because if they are too many, the entire data is rejected. In situations where they cannot be ignored, for example, where the dataset is limited, small, or where recordings from subjects cannot be prevented. The best solution is to remove those artifacts and restore the signal. Different approaches are used to remove motion artifacts, like recording additional data on the subject's movement using referenced channels. [5], [6] and [7].

#### **2.1.1.2 Instrumental Artifacts**

Spikes can originate from a momentary change in impedance at a given electrode. Using basic low-pass filtering methods, instrumental noise is a random noise that can be removed. After data conversion to the frequency domain, a method such as Moving Average and cutting off higher frequencies are used. The sensitivity of these methods needs to be manually determined to avoid data distortion [5],[8].

#### **2.1.1.3 Physiological Artifacts**

Sweating or changes in temperature may cause electrode drifts so when coping with noise sources for physiological use, potentials of skin's outer layer, and ionic features it is important to regard the capacity of sweat glands also. The solution to handle this problem is to reduce the skin potential and improve the signal to noise ratio, by using abrasive cream on the skin. [8] and [9].

#### **2.1.1.4 Eye Blink Artifacts**

The large amplitude deflection in frontal region is Eye Blink Artifacts, these are maximal at FP1 and FP2 because eye balls are closer to frontal region. [10]

Power line noise can be removed using a notch filter. For removing the noise from the raw dataset band pass filter can be used. The lower cut-off frequency used in this dataset is 0.5 Hz and a higher cut-off frequency at 100 Hz.

### **2.1.2 Transcranial Direct Current Stimulation (tDCS)**

Transcranial direct current stimulation (tDCS) is a non-invasive brain stimulation technique that consists of applying a constant, low electric current between electrodes over the scalp in order to modulate synaptic activity [11]. Anodal stimulation, considered an excitatory stimulation, has been shown to lead to the depolarization of the resting membrane potential, reducing the threshold required for neuronal firing and increasing neuronal excitability [12]. On the other hand, cathodal or inhibitory stimulation, has led to hyperpolarization of the resting membrane potential and decreased neuronal excitability [13]. tDCS is a form of neuro stimulation that modulates the membrane potential of neurons in the cerebral cortex by a low-intensity direct current (1-2 mA). The electrical current is applied directly to the scalp at the targeted brain area (i.e., primary motor cortex, prefrontal dorsolateral, etc.) through sponge electrodes soaked with saline solution or rubber electrodes with conductive gel. It is a low-cost technique, with easy application and practically no adverse effects [14].

tDCS is a low-cost technique with minimal adverse effects and easy application. It has also been proposed that the effects of tDCS are network-activity dependent, requiring active neurons to modulate upon [15]. It is a safe method for treating various therapeutic option for various conditions, such as neuropsychiatric disorders pain syndromes rehabilitation and a tool for modulating cortical activity. Based on the ability of tDCS to modulate synaptic transmission, researchers have investigated whether tDCS can improve mood and cognitive symptoms. Results from clinical trials administering tDCS have demonstrated improved working memory performance and enhanced episodic memory in a variety of populations, including Parkinson's disease [16], patients with depression and healthy individuals.

There have been a handful of studies that have investigated the effects of tDCS on cognitive performance [17]. Studies have shown that repeated sessions of tDCS significantly improved global cognition in mild Alzheimer disease (AD) patients. In addition, working memory, recall, and frequently recognition memory has improved following single and repeated sessions of tDCS in mild AD [17]. Based on recent interest in the activity-dependent model, which suggests that tDCS effects may be greater when applied to active neurons [17], studies have begun to investigate the use of combining other cognitive enhancing interventions with tDCS to prime neurons of interest. In healthy adults, tDCS applied during a cognitive task resulted in greater improvement in working memory performance compared to when tDCS was applied at rest and when sham tDCS was applied during the task [13]. In the AD

population, there have been only a few studies that have investigated combination therapies with tDCS, and all of them have used cognitive training to enhance cognitive skills through repetitive tasks or activities related to memory, attention, or other cognitive functions [17]. A small number of studies have reported that tDCS with cognitive training was associated with greater improvements on the digit span and trained and untrained picture-naming tasks, compared to sham tDCS with cognitive training [10]. tDCS has also been previously used in numerous studies involving older, frail patients with no serious adverse events noted [14]. In a review of 117 tDCS studies, commonly reported side-effects included itching, tingling, headache, discomfort, and a burning sensation. The prevalence of those side-effects was not significantly different between active and sham groups.

## **2.2 EEG and tDCS for Brain Rehabilitation**

A review is presented on the past work where the tDCS therapy is used for brain rehabilitation in conjunction with different brain imaging techniques such as EEG and functional near infrared spectroscopy (FNIRS).

Over the years different studies have proved the efficacy of tDCS in reducing depressive symptoms and improving cognitive functioning of depressed patients. The neurophysiological mechanisms involved in the antidepressant effects of tDCS remain incompletely understood [10]. Powell studied the neuro modulatory effects of tDCS on cortical activity for the treatment of mood disorders using EEG and he reveals an asymmetry in EEG frontal alpha activity, i.e. lower alpha power in the right hemisphere compared to the left in subjects having depression [10]. He further used EEG to study the modulatory effect of tDCS on changes in cortical activity in subjects with mood disorders.

Wozniak-Kwasniewska [18] showed that EEG oscillatory activity was significantly different for depressed patients that responded to repetitive transcranial magnetic stimulation (rTMS) therapy compared to non-responders, suggesting that baseline EEG has predictive value for brain stimulation treatment outcomes.

Alaa M. Al-Kaysia, [19] investigate the feasibility of identifying major depressive disorder (MDD) patients that respond to tDCS treatment based on resting-state electroencephalography (EEG) recorded prior to treatment commencing and machine learning techniques. Their findings demonstrate the feasibility to identify patients that will respond to tDCS treatment [13].

In this study, we sought to identify features of EEG recorded at baseline, during the course of stimulation and after the completion of stimulation therapy to analyze the efficacy of tDCS therapy. In this study we used Event related De synchronization (ERDs) to predict the improvement in amplitude and cognition following tDCS treatment based on spectral power of EEG. We further used Common spatial pattern (CSP) to analyze the strength of EEG signals after stimulation therapy with before the therapy.

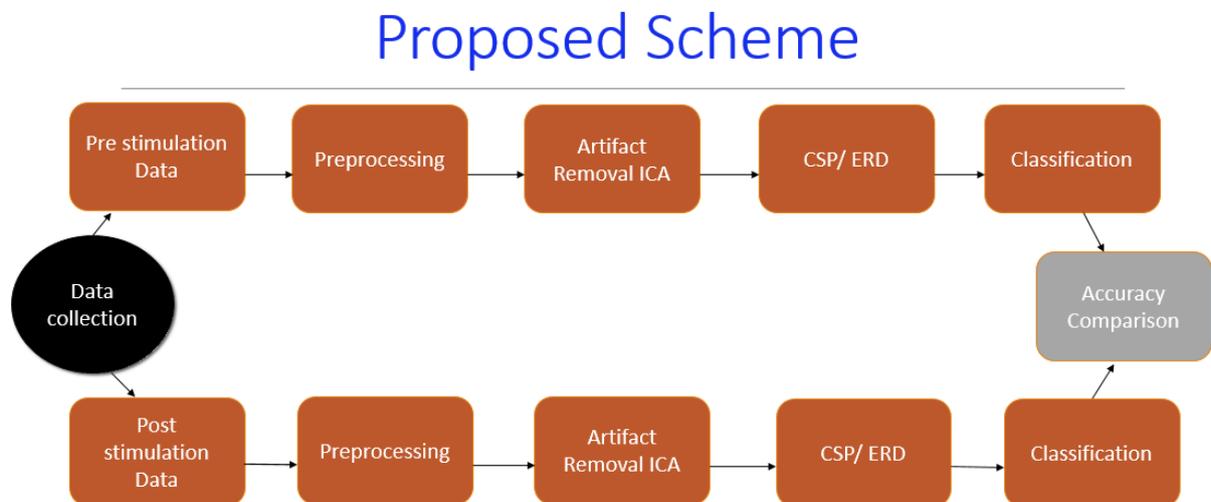
## CHAPTER 3: METHODOLOGY

This chapter explains the general experimental design for brain-cognitive state improvement estimation using tDCS and EEG measurement and the datasets used in this work. Two datasets are used in this work, Dataset one is recorded in the lab using our own experimental design and setup. Dataset two used is open access data recorded by Department of Electrical, Electronic and Information Engineering (DEI), University of Bologna, Cesena, Italy and is available online at Dryad Digital Repository. Raw EEG data can be accessed from the Dryad Digital Repository at:

<https://doi.org/10.5061/dryad.3m8j0>.

### 3.1 Proposed Scheme

A scheme is proposed in this study for which data is collected. For each of the collected data set used it is first group into two pre stimulation data and post stimulation data then preprocessed, ICA is applied to remove artifact and at the end after pattern recognition techniques Classification algorithms are applied to them and their accuracies are compared.



**Figure 3-1:** Proposed Scheme

### 3.2 Data Collection

In this study two datasets are collected. Dataset 1 is recorded in the lab whereas dataset 2 used is open-source data. The results of classification accuracies are compared by applying the technique on both datasets and cognitive state improvement is estimated.

### 3.2.1 Dataset 1 Experimental Paradigm

The study is performed in lab with subject sitting in front of shared screen, wearing EEG open BCI headset, Brain driver tDCS device set to his head at the start of the session. The signals from the EEG head sets were recorded in the laptop via Bluetooth while open BCI software running on the laptop during the study. For each subject during the experimental session video is recorded via the webcam of the screen and also through external android camera. Figure 3.1 and 3.2 shows experimental setup and recording of subject.

Total of 10 healthy volunteers all male participated in the study, aged between 24 and 35 medians 27. All participants signed a written form to participate in the study. The study was conformed and conducted according to NUST Ethics committee, During the study blood pressure and temperature of the participant were monitored and recorded in the table 3-1. During the whole study and after completion none of the participant show signs of discomfort such as headache and pinching.

**Table 3-1:** Participants of the study

<b>Sr No</b>	<b>Subject Name</b>	<b>Before Experiment BP</b>	<b>After Experiment BP</b>	<b>Temp F</b>	<b>Age</b>
<b>1</b>	<b>Kalim Ullah</b>	79/120	94/130	98.5	35
<b>2</b>	<b>Abdul Basit</b>	80/120	90/118	98.7	27
<b>3</b>	<b>Talha Rashid</b>	76/123	76/117	98.3	26
<b>4</b>	<b>Manzoor Ahmed</b>	79/127	79/133	98.6	27
<b>5</b>	<b>Ahmad Zubair</b>	74/114	78/108	98.1	26
<b>6</b>	<b>Abdul Haseeb</b>	76/123	75/126	99.2	24
<b>7</b>	<b>Areej Khan</b>	63/106	63/106	98.9	33
<b>8</b>	<b>Asghar Ali</b>	77/120	73/114	98.2	33
<b>9</b>	<b>Naveed Hussain</b>	78/131	76/124	98.4	26
<b>10</b>	<b>Mudassar Ayub</b>	80/132	83/131	98.9	29

All Participants received Anodal stimulation for total of 300 seconds with Anode placed over **Cz** and cathode at **Cp1**. tDCS was delivered through The Brain Driver tDCS v 2.1 and the intensity of stimulation was set to 2 mA.

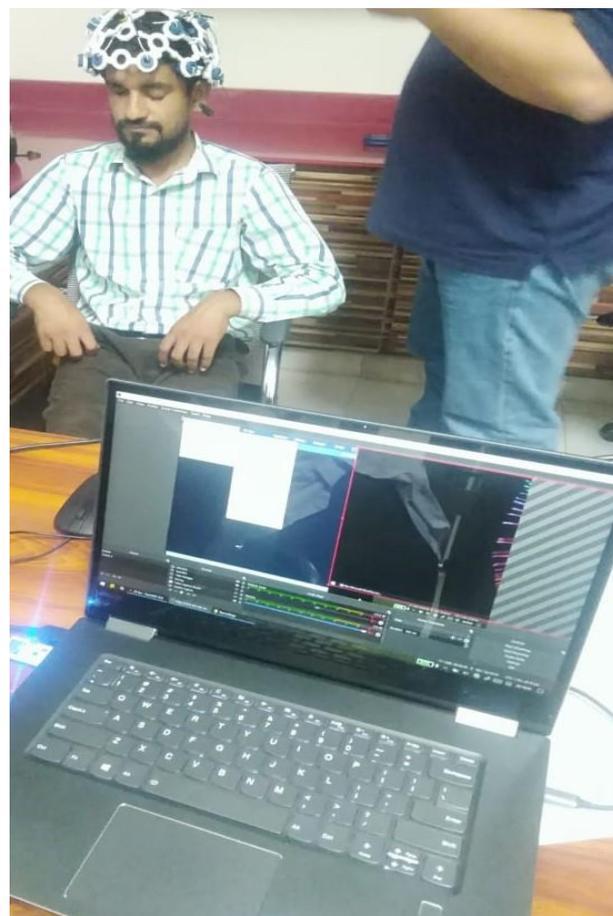
EEG Signals are recorded with Open BCI Ultracortex Mark IV EEG headset. The Ultracortex Mark IV is capable of sampling up to 16 channels of EEG from up to 35 different 10-20 locations. In the current study we use 16 channels at sampling frequency of 128 Hz. EEG channels for recording were placed in both right and left hemisphere. Table 3-2 shows the montage, channel number and channel name of different EEG channels used

**Table 3-2:** Recording Montage.

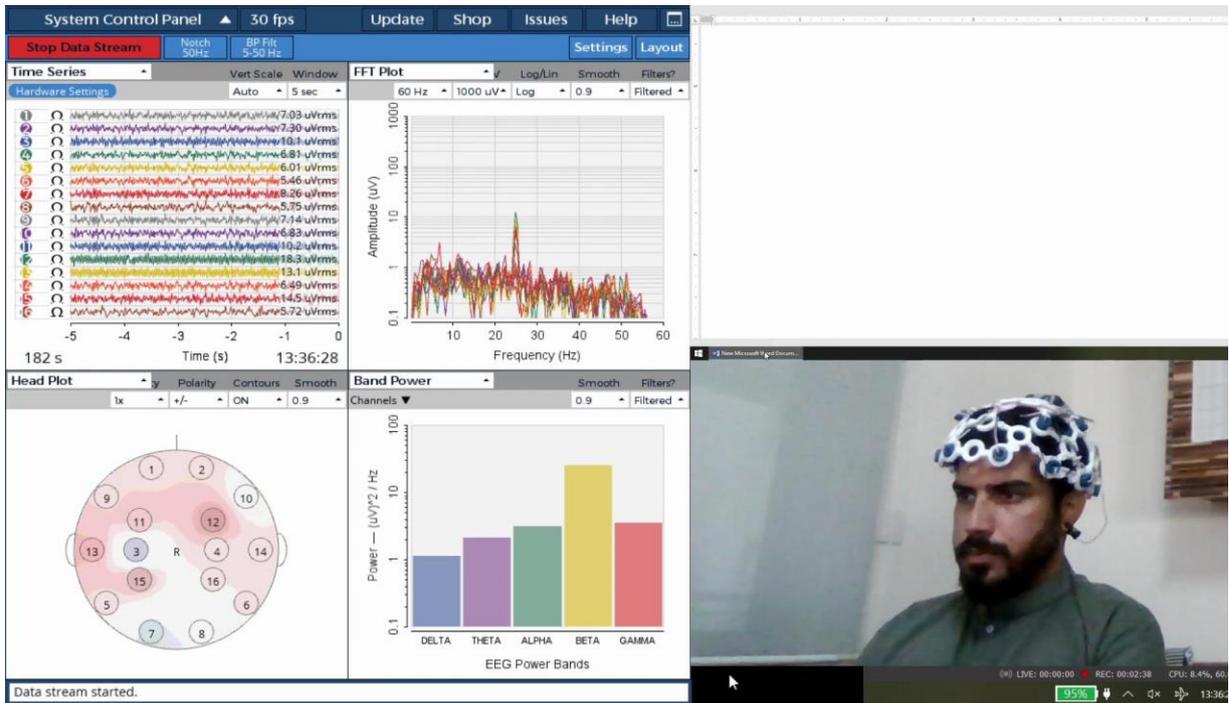
<b>Channel number on montage</b>	<b>Channel Name</b>
1	Fp1
2	Fp2
9	F7
11	F3
12	F4
10	F8
13	T3
3	C3
4	C4
14	T4
5	T5
15	P3
16	P4
6	T6
7	O1
8	O2

EEG signals are recorded for total of 9 minutes for each subject in 3 equal intervals. Interval 1 is 3 minutes EEG recording before tDCS stimulation. Interval 2 is 3 minutes during stimulation and interval 3 is 3 minutes recording after tDCS stimulation. All 3 intervals are recorded in one go without any delay between them.

Further each interval is divided into Motor Imagery (MI) Task period and Rest Period or Non MI period. During each 300 Seconds of recording interval there are five Rest periods of approximately 20 seconds and five Motor Imagery periods of 40 Seconds. The distribution of these tasks in each session is random. When white screen is presented to the subjects Motor Imagery period time starts during which subjects are instructed to perform Motor imagery tasks in their mind such as limb motion, running on the screen, playing soccer, arms movement and walking etc. while during the rest period when white screen is removed from the LED screen Rest period starts during which subjects are instructed to relax their mind without thinking about any activity which involves motion. Table 3-1 to 3-3 shows distribution of MI and Rest period tasks for one of the subject for all three intervals



**Figure 3-2:** Recording Setup



**Figure 3-3:** Subject Recording during Experimental Setup

**Table 3-3:** Pre-Stimulation Motor Imagery Period (MIP) & Rest Period (RP)

Interval Name	Interval Time	Start	Interval End Time	Interval Duration (s)
<b>Pre-Stimulation</b>		'13:33:50'	'13:39:00'	<b>310</b>
<b>Pre MIP/RP No</b>	<b>Start Time</b>		<b>End Time</b>	<b>Duration</b>
<b>Pre MIP 1</b>		'13:34:02'	'13:35:00'	58
<b>Pre RP 1</b>		'13:35:01'	'13:35:20'	19
<b>Pre MI P 2</b>		'13:35:21'	'13:36:00'	39
<b>Pre RP 2</b>		'13:36:01'	'13:36:19'	18
<b>Pre MIP 3</b>		'13:36:20',	'13:37:00'	40
<b>Pre RP 3</b>		'13:37:01'	'13:37:19'	18
<b>Pre MIP 4</b>		'13:37:20'	'13:38:00'	40

<b>Pre RP 4</b>	'13:38:01'	'13:38:14'	13
<b>Pre MIP 5</b>	'13:38:15'	'13:38:51'	36
<b>Pre RP 5</b>	'13:38:52'	'13:39:00'	8

In video Pre MI time start at '13:33:50' but in text file recording stats at '13:33:26' so it is required to verify in code when data is taken in code

**Table 3-4:** During Stimulation Motor Imagery Period (MIP) & Rest Period (RP)

<b>Interval Name</b>	<b>Interval Start Time</b>	<b>Interval End Time</b>	<b>Interval Duration</b>
<b>During Stimulation</b>	'13:39:01'	'13:44:00'	<b>299</b>
<b>During MIP/RP No</b>	<b>Start Time</b>	<b>End Time</b>	<b>Duration</b>
<b>During MIP 1</b>	'13:39:20'	'13:40:00'	40
<b>During RP 1</b>	'13:40:01'	'13:40:19'	18
<b>During MIP 2</b>	'13:40:20'	'13:41:00'	40
<b>During RP 2</b>	'13:41:01'	'13:41:19'	18
<b>During MIP 3</b>	'13:41:20'	'13:42:00'	40
<b>During RP 3</b>	'13:42:01'	'13:42:20'	19
<b>During MIP 4</b>	'13:42:21'	'13:43:00'	39
<b>During RP 4</b>	'13:43:01'	'13:43:14'	13
<b>During MI P5</b>	13:43:15'	'13:43:52'	37
<b>During RP 5</b>	'13:43:53'	'13:44:00'	7

**Table 3-5:** Post Stimulation Motor Imagery Period (MIP) & Rest Period (RP)

Interval Name	Interval Start Time	Interval End Time	Interval Duration
Post Stimulation	'13:44:01'	'13:49:01'	<b>300</b>
Post MIP/RP No	Start Time	End Time	Duration
Post MIP 1	'13:44:20'	'13:45:00'	40
Post RP 1	'13:45:01'	'13:45:19'	18
Post MI P 2	'13:45:20'	'13:46:00'	40
<b>Post RP 2</b>	'13:46:01'	'13:46:19'	18
<b>Post MI P 3</b>	'13:46:20'	'13:47:01'	41
<b>Post RP 3</b>	'13:47:02'	'13:47:19'	17
<b>Post MI P 4</b>	'13:47:20'	'13:47:59'	39
<b>Post RP 4</b>	'13:48:00'	'13:48:14'	14
<b>Post MI P 5</b>	'13:48:15'	'13:48:50'	35
<b>Post RP 5</b>	'13:48:51'	'13:49:01'	10

Motor imaginary tasks are performed randomly during each interval. Its mapping to the dataset (recorded EEG text file with one column as timestamp) will be performed through the recorded video.

### 3.2.2 Dataset 2 Experimental Paradigm

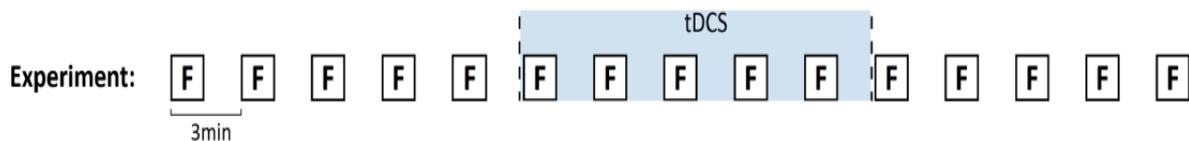
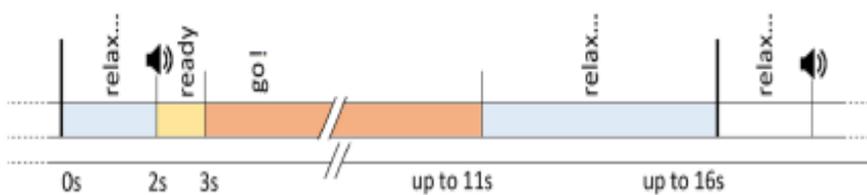
The dataset used is published at DRYAD directory and is contributed by Mondini, Valeria, Mangia, Anna Lisa and Cappello, Angelo from University of Bologna [2]

The dataset is organized such that the EEG data collected from the twenty healthy volunteers participating in the study are firstly grouped according to the type of stimulation they received (anodal or cathodal stimulation) [2].

Inside each stimulation folder, data are grouped according to the subject (S01, S02... to S10). For each subject, data are grouped according to the experimental day (day 1 and day2), one corresponding to real and one to sham stimulation.

Each day was composed by 15 runs of BCI operation [2] with feedback (neurofeedback01 to neurofeedback15). Runs 01-05 correspond to the condition “before”, runs 06-10 to the condition “during” and runs 11-15 to the condition “after” stimulation. The dataset also includes the calibration runs (calibration01, calibration02...), collected at the beginning of the first experimental day 1.

### Feedback trial:



**Figure 3-4:** Experimental Setup

Each “neurofeedback” or “calibration” folder finally contains a “.txt file”, which includes:

- EEG raw data collected in the corresponding run from the 12 EEG electrodes used in the study
- Reference signal from right ear lobe
- Additional column indicating the trial condition (rest, ready, motor imagery...)

Data are collected with sampling frequency of  $F_s=128\text{Hz}$  and are stored in the “.txt” files as arrays, where rows are samples and columns are channels. The first 12 columns contain the 12 EEG channels, in the following order:

[Fc<sub>z</sub> Fc<sub>2</sub> Fc<sub>4</sub> Fc<sub>6</sub> Cz C<sub>2</sub> C<sub>4</sub> C<sub>6</sub> Cp<sub>z</sub> Cp<sub>2</sub> Cp<sub>4</sub> Cp<sub>6</sub>]

Recording is done with the ground electrode in Pz. The 13th column contains the reference signal from right ear lobe. The 14th column contains the information on the trial condition, notably:

- “0” \_ for samples in the “rest” period before the warning tone (lasting 2s).
- “1” \_ for samples in the “ready” period after the warning sound and before motor imagery (lasting 1s).
- “2” \_ for samples in the “motor imagery” period (4s during calibration, or up to 8s during neuro feedback).
- “3” \_ for samples in the “rest” period concluding the trial (lasting 5s).

### **3.3 Data Preprocessing:**

For dataset 1 each subject file has 15 sessions of recording n1 to n5 represents pre stimulation data n6 to n10 is during stimulation data whereas n11 to n15 session files are for post stimulation data. Raw anodal data is first read with all 14 channels included. Runs 6 to 10 which are the EEG data recording during the tDCS stimulations are excluded in the read raw data. The data is read as numpy array with shape ( number of runs, number of channels, number of EEG samples in each run).Sampling frequency is 128 Hz and FIR filter is applied with lower pass band edge 8 Hz and higher pass band edge of 22 Hz .Channel 13 which contains reference signal from the right ear lobe

#### **3.3.1 Data Labeling**

Raw data is recorded using Open BCI setup as text file with total of 22 columns. Column 1 to 16 contains the raw EEG recorded signals data, columns 17 to 19 records the accelerometer data while column 20 and 21 records the time stamp data the first one records the standard time that appears on the computer while the later column records the absolute time of computer (CPU) clock respectively. Two columns are added at the end to label the events in the data. Column 22 is added with header stim where 0 in the column represents data recorded before the tDCS stimulation is applied to the subject, 1 shows data recorded while stimulation is applied to the subject and 2 labels the data after the stimulation ends. The last column added to the data with header MI has two values 0 and 1. Number 0 where appears in the data shows that the data is recorded while the subject is at Relax and shows Rest Period (RP), 1 in this column represents the subject is performing Motor Imagery Period (MIP) task while the data is being recorded.

### 3.3.2 Intra Correlation

There are total of 3 segments in the data, Pre stimulation, during stimulation and Post stimulation. Further each of the segment contains two tasks that is instructed to each subject to perform, Motor Imagery Period (MIP) Task and Rest Period/Non MI (RP) period tasks. Linear correlation among these different tasks period in each segment is performed through df.corr() function in python. As the correlation requires data frame so data frame from the initial data is generated using pd.DataFrame() function.

For each of the MI Task period we select total of 10 seconds of patch of data with 5 second delay from the **Interval Start time** of that particular MI Task Period to settle any transition in the data if so e.g. Pre MI 1 task period will start from '13:34:07' instead of '13:34:02' and will last up to '13:34:17'. We take 5 MI task periods in a correlation Matrix. Similarly, for Rest periods tasks we set a delay of 3 seconds from the time where the corresponding prior MIP period ends that is prior MIP **Interval End time** so RP1 starts at '13:35:03' and last for '13:35:13' total of 10 seconds and these settings of MIP and RP Tasks holds true for all 3 data segments periods Pre, during and Post.

Total of 5 MIP tasks and 4 corresponding RP tasks are used to calculate the intra correlation matrix of order 9\*9. The correlation coefficient value of 0 in the correlation matrix represents no relation of row variable with the column variable for that particular cell where Correlation coefficient value of 1 represents high correlation. The correlation of a variable with itself is 1. For that reason, all the diagonal values are 1.00. Below is the intra correlation matrix of Motor Imagery Period (MIP) tasks and Rest Period (RP) tasks of **Pre stimulation time Segment**.

**Table 3-6:** Pre stimulation time MI and RP Intra Correlation Matrix

	<b>MIP1</b>	<b>MIP2</b>	<b>MIP3</b>	<b>MIP4</b>	<b>MIP5</b>	<b>RP1</b>	<b>RP2</b>
<b>MIP1</b>	1.000000	0.971274	0.924716	0.843433	0.944520	0.283616	0.141596
<b>MIP2</b>	0.971274	1.000000	0.973997	0.917610	0.762981	0.191188	0.082747
<b>MIP3</b>	0.924716	0.973997	1.000000	0.964014	0.848487	0.169095	0.091957
<b>MIP4</b>	0.843433	0.917610	0.964014	1.000000	0.935816	0.196109	0.137582
<b>MIP5</b>	0.944520	0.762981	0.848487	0.935816	1.000000	0.147425	0.193461
<b>RP1</b>	0.283616	0.191188	0.169095	0.196109	0.147425	1.000000	0.980007
<b>RP2</b>	0.141596	0.082747	0.091957	0.137582	0.193461	0.980007	1.000000

**RP3** 0.090278 0.153307 0.176535 0.180192 0.174317 0.923153 0.960869  
**RP4** 0.115201 0.103414 0.187447 0.153663 0.088663 0.791279 0.836247

	<b>RP3</b>	<b>RP4</b>
<b>MIP 1</b>	0.090278	0.115201
<b>MIP 2</b>	0.153307	0.103414
<b>MIP 3</b>	0.176535	0.187447
<b>MIP 4</b>	0.180192	0.153663
<b>MIP 5</b>	0.174317	0.088663
<b>RP1</b>	0.923153	0.791279
<b>RP2</b>	0.960869	0.836247
<b>RP3</b>	1.000000	0.904568
<b>RP4</b>	0.904568	1.000000

**Table 3-7:** Post stimulation time MI and RP Intra Correlation Matrix

	<b>MIP1</b>	<b>MIP2</b>	<b>MIP3</b>	<b>MIP4</b>	<b>MIP5</b>	<b>RP1</b>	<b>RP2</b>
<b>MIP1</b>	1.000000	0.707674	0.816236	0.800155	-0.807727	0.199718	0.172314
<b>MIP2</b>	0.707674	1.000000	0.810242	0.862333	0.877322	0.116695	0.181079
<b>MIP3</b>	0.816236	0.810242	1.000000	0.890088	0.799490	0.169125	0.112501
<b>MIP4</b>	-0.800155	0.862333	0.890088	1.000000	0.968643	0.307999	0.807234
<b>MIP5</b>	-0.807727	0.877322	0.799490	0.968643	1.000000	0.128477	0.173318
<b>RP1</b>	0.199718	0.116695	0.169125	0.307999	0.128477	1.000000	0.696833
<b>RP2</b>	0.172314	0.181079	0.112501	0.807234	0.173318	0.696833	1.000000
<b>RP3</b>	0.005949	0.146368	0.139793	0.181639	0.133164	0.833837	0.847200
<b>RP4</b>	-0.179609	0.231322	0.011913	0.148858	0.155096	0.868145	0.731096

	<b>RP3</b>	<b>RP4</b>
<b>MIP1</b>	0.005949	-0.179609
<b>MIP2</b>	0.146368	0.231322
<b>MIP3</b>	0.139793	0.011913

MIP4 0.181639 0.148858  
MIP5 0.133164 0.155096  
RP1 0.833837 0.868145  
RP2 0.847200 0.731096  
RP3 1.000000 0.901991  
RP4 0.901991 1.000000

### 3.3.3 Data Standardization

Standardization transforms data to have a mean of zero and a standard deviation of 1. Data is standardized using the below formula. This standardization is called a **Z-score**.

$$x_{new} = \frac{x - \mu}{\sigma}$$

where x is the data point

$\mu$  is the data mean

$\sigma$  is standard deviation

The idea is to allow different data sets to be comparable. Once you compute the mean, you then want to see how the data varies about the mean. The purpose of subtracting the mean from a dataset is to obtain a dataset whose mean is zero. Dividing by the standard deviation lets you compare the data distribution with a normal distribution.

### 3.4 Data Filter

The original shape of data is (116933, 22) the data is first converted to numpy array and 16 channels of interest that contains EEG data are selected and their transpose is taken such that the data shape becomes (16, 116933). To apply filter and observe PSD, raw object is created using `mne.io.RawArray()` in python, this function takes data array as parameter and info as attribute. The attribute info contains information such as channel names, channel types and sampling frequency (128 Hz). Raw object data is then filter using Band Pass Filter with lower cut-off frequency ( $f_L$ ) of 1 Hz and higher cut-off frequency ( $f_H$ ) of 40 Hz.

### 3.4.1 ICA

Independent Component Analysis (ICA) tries to maximize independence by finding linear transformation of feature space (observed data variables) to a new feature space (hidden/source data variables) such that each of the new individual features are statistically mutually independent i.e. the new features mutual information is zero, whereas the mutual information of all new features in original feature space is as high as possible. ICA reconstruct data by predicting latent/source features from observable features and vice versa.

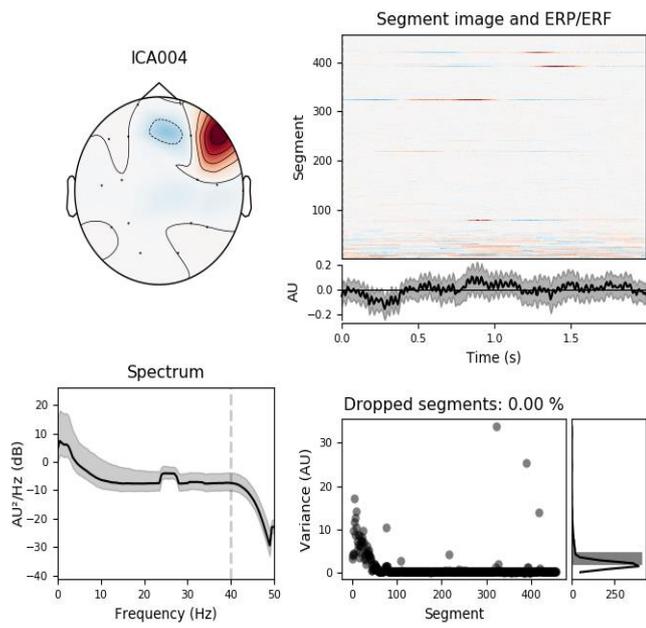
ICA defines a generative model for the observed multivariate data, which is typically given as a large database of samples. In the model, the data variables are assumed to be linear mixtures of some unknown latent variables, and the mixing system is also unknown. The latent variables are assumed Non Gaussian and mutually independent, and they are called the independent components of the observed data. These independent components, also called sources or factors, can be found by ICA.

Typical examples are mixtures of simultaneous speech signals that have been picked up by several microphones, brain waves recorded by multiple sensors, interfering radio signals arriving at a mobile phone, or parallel time series obtained from some industrial process.

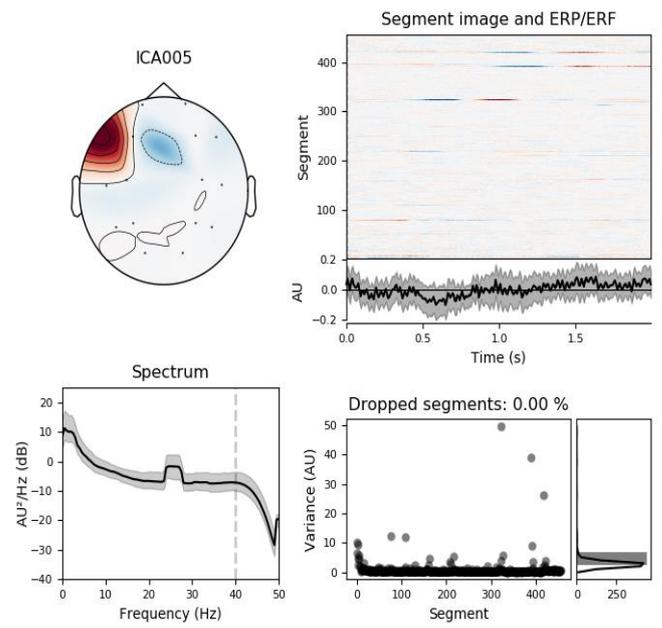
Many EEG signals including biological artifacts reflect non-Gaussian processes. MNE-Python supports identifying artifacts and latent components using temporal ICA. MNE-Python implements the **mne.preprocessing.ICA** () class that facilitates applying ICA to EEG data. There are 16 recorded channels for EEG data so we use 16 components for ICA. Figure 3.5 shows all the ICA components with heat maps, segmented ERP and spectrum



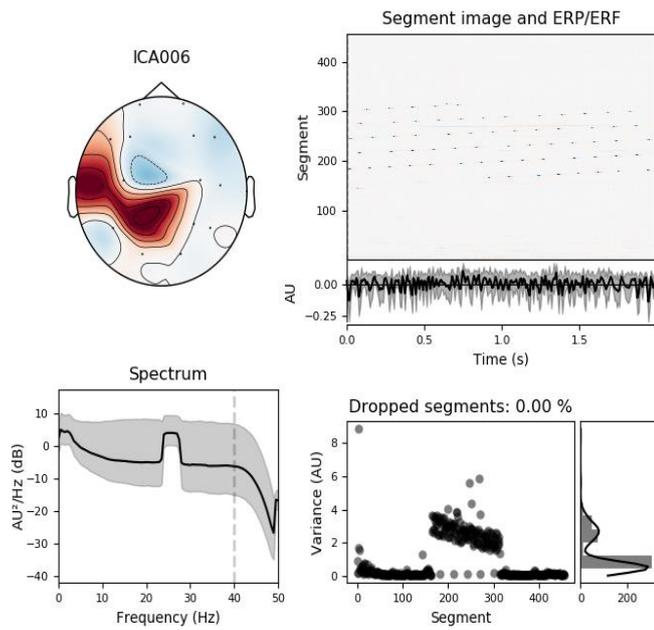
### ICA 4



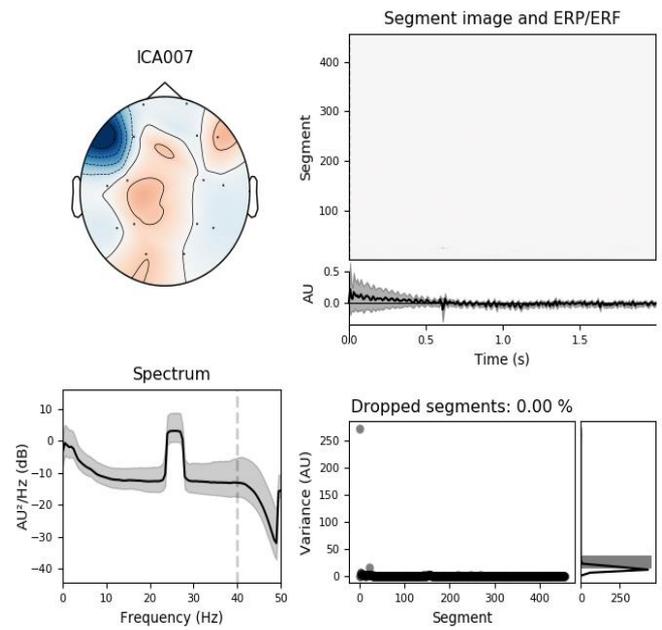
### ICA 5



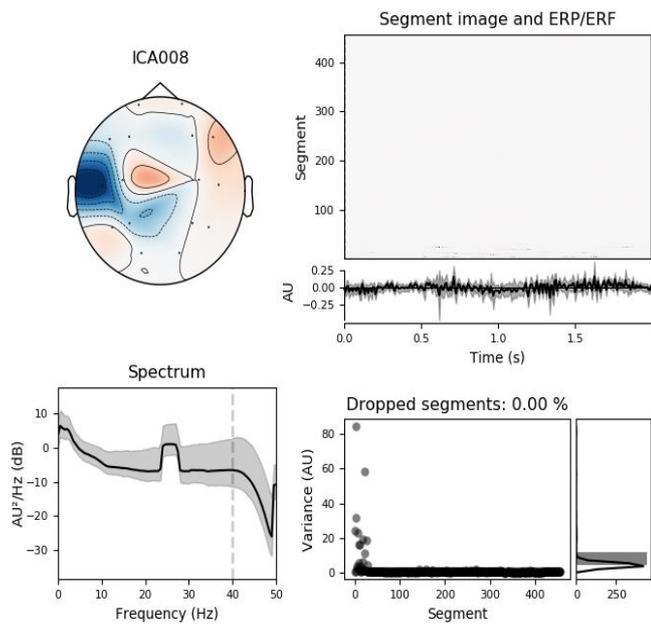
### ICA 6



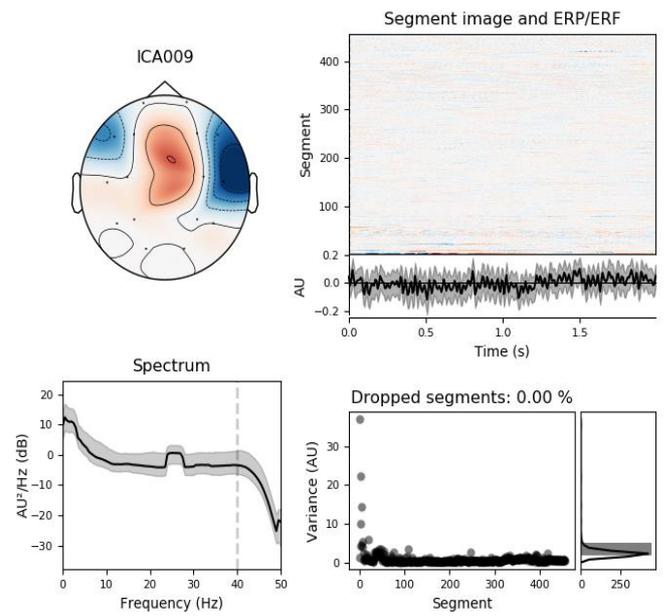
### ICA 7



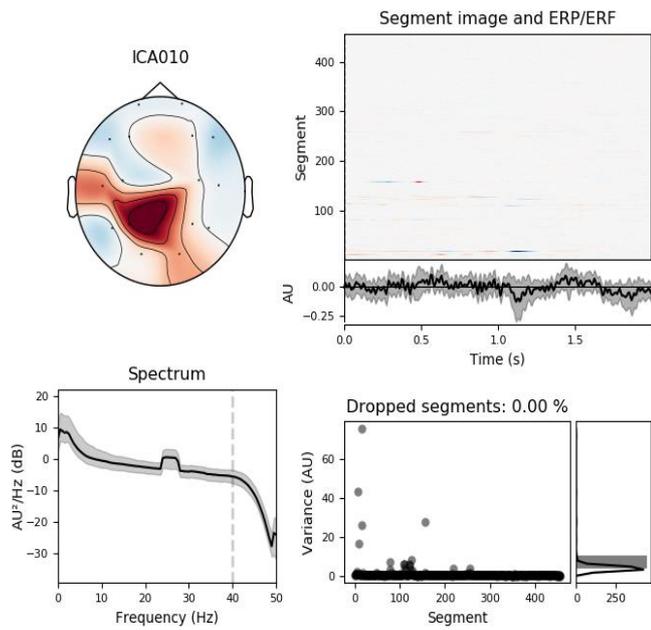
### ICA 8



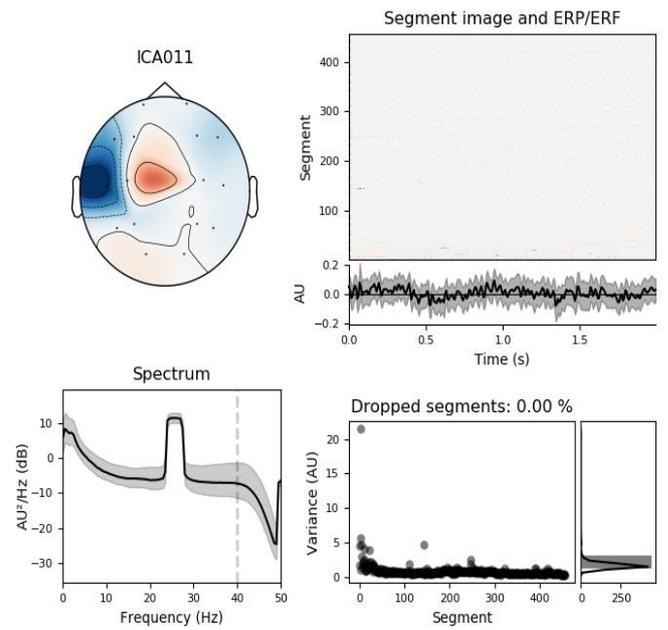
### ICA 9



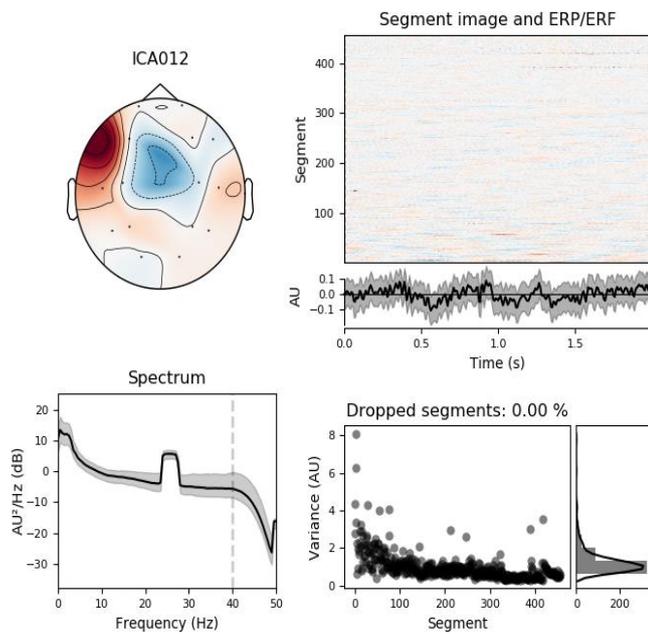
### ICA 10



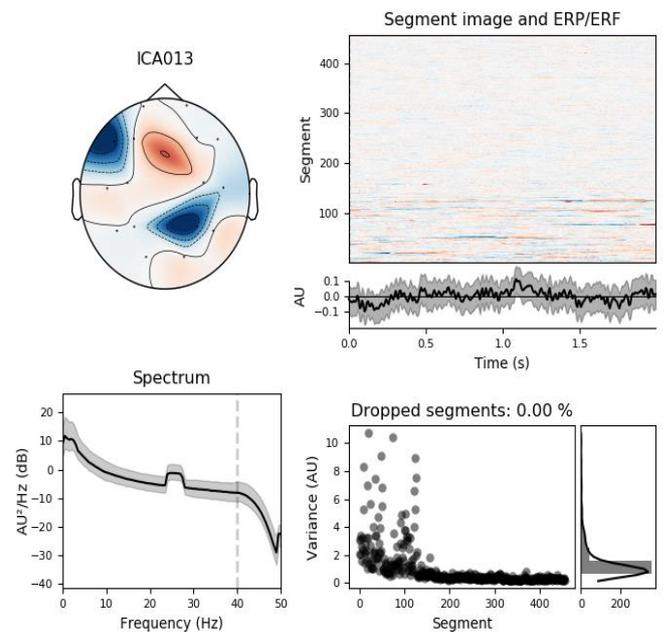
### ICA 11



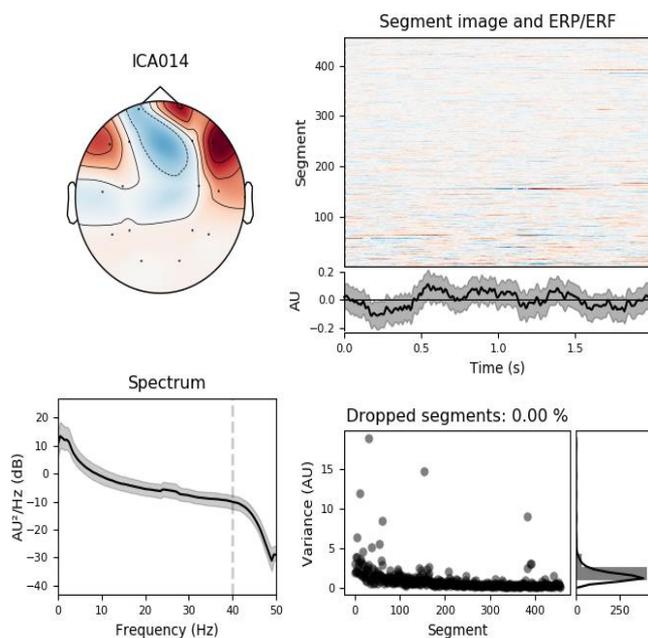
## ICA 12



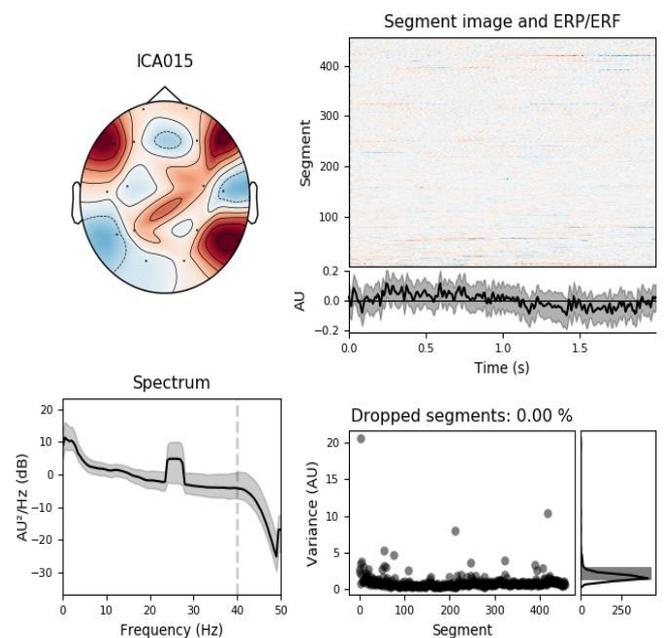
## ICA 13



## ICA 14



## ICA 15

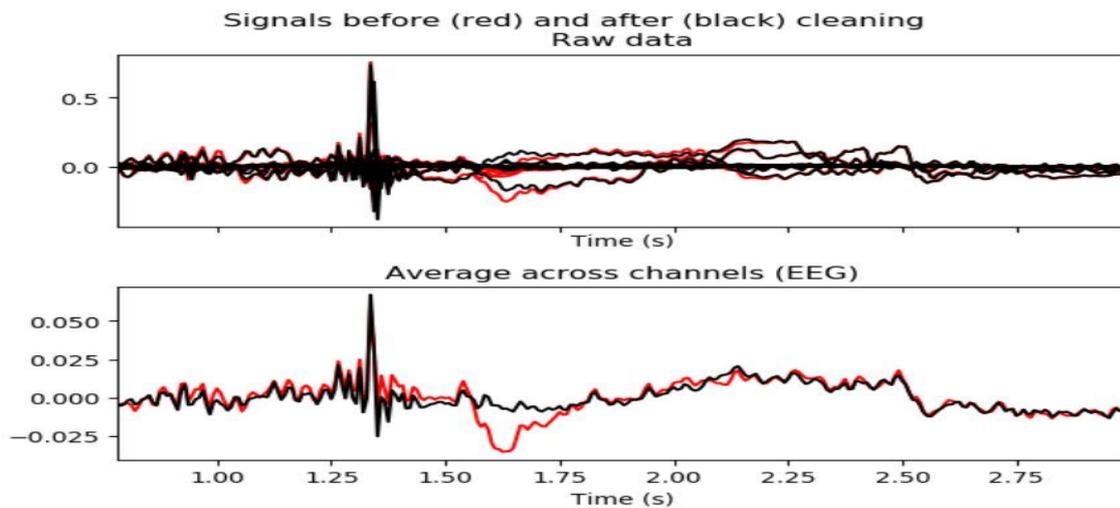


**Figure 3-5:** ICA components with heat maps, segmented ERP and spectrum

### 3.4.2 ICA components Rejection

Bad data segments can be excluded from the model fitting by reject parameter. Raw data is clean by removing ICA components that shows bad data segments such as sudden drift in amplitude, spikes etc. while manually observing the ICA components. Data is cleaned by excluding the ICA components 1, 3, 6 and 11 from the data.

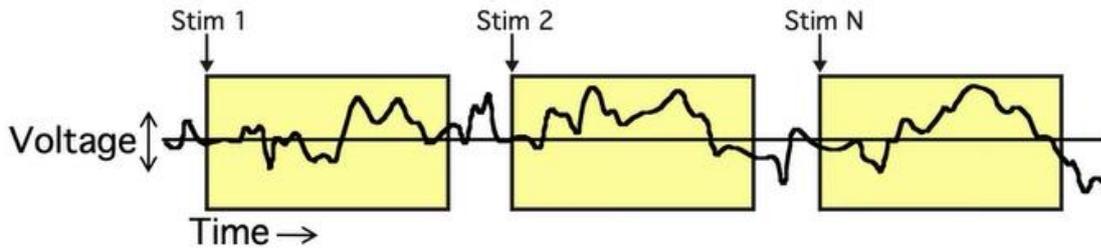
Figure 3.6 shows signal before and after cleaning the data through ICA.



**Figure 3-6:** The signal before and after cleaning the data through ICA

### **3.5 Event Related Potential (ERP)/Event Related Desynchronization (ERD)**

ERP is the electrical potentials (Voltages) that are related to specific events. After observing the raw EEG there isn't very consistence response following each stimulus. This is because brain is doing millions of other things in addition to processing the little stimulus. To pullout the brain response to the stimulus a simple procedure is used called Signal Averaging. First we grab the period of EEG following each stimulus this is called segmenting the EEG. Each segment is called Epoch.



**Figure 3-7:** EEG signal with stimulus onset

Now we can take all EEG Epochs and line them up in time. The epoch length depends on the nature of specific experiment, but usually it's between 10 milliseconds to 2000 milliseconds. To pullout the brains consistent response to some type of event we can simply average across the epochs for that event type. When we average across enough epochs any activity that is consistent from trial to trial remains in the average, and any random noise simply averages out. This gives us an averaged ERP waveform. Whereas Event-related DE synchronization (ERD) is the phasic relative power decrease of a certain frequency band occurring in relation to stimulation. ERD are negative values relative to the baseline. The value of "2" for samples in the EEG data in column 14 of data set represents motor imagery "MI" period which is up to 8s induration of neuro feedback runs.

EEG segment of 0 to 8 Seconds from the prepare cue was then extracted to perform offline ERD analysis. Following method was used to compute the ERD strength values [20] for all channels in pair of 4 during a single run.

- Band pass filtering of 8 to 22 Hz on the EEG time segment - 3 to 8 seconds relative to prepare cue for all trials.
- Squaring the band pass-filtered samples to obtain power samples.
- Average power samples across all trials. Compute power of baseline from average on time segment -3 to 0 seconds.
- Compute ERD/ERS strength values of channel j on time segment 0 to 8 seconds using the following equation:

$$S_j(t) = \frac{A_j(t) - R_j}{R_j} \times 100$$

where  $A_j(t)$  is the averaged power sample of time sample t of channel j and  $R_j$  is the averaged power of baseline of channel j .

- Compute the ERD strength value from the sum of the negative values for time samples  $t$  from 0 to 8 seconds of channel  $j$  using the following equation:

$$E_j = \sum_{t \in [2,6]} (S_j(t) | S_j(t) < 0)$$

The ERD/ERS value was calculated by averaging the absolute power according to different range of frequency bands. The outcome ERD plots were drawn from  $-3$  to  $8$  s total of  $11$  seconds. The baseline period was from  $-3$  to  $0$  s. The frequency bands of interest were alpha band ( $8-12$  Hz), low beta band ( $13-20$ ) overall  $8$  to  $24$  Hz. The maximum and minimum value of event related power in the ERD plots are set to  $1.5$  and  $-1$  respectively.

There is no delay between when the brain activity happens and when the voltage is picked up by our scalp electrodes, so a voltage at  $293$  ms reflects brain activity that happened exactly at  $293$  ms. ERP allows to see the flow of information through the brain millisecond by millisecond.

### 3.6 Common Spatial Patterns (CSP)

The common spatial patterns (CSP) algorithm is a feature extraction method that uses spatial filters to maximize the discriminability of two classes.

The method of common spatial patterns (CSP) designs spatial filters in such a way that the variances in the filtered time series data are optimal (in the least squares sense) for discrimination. CSP is denoted by

$$S = W^T E \quad \text{or} \quad s(t) = W^T e(t),$$

Where  $W$  is spatial filter matrix,  $S$  is filtered signal matrix.

For CSP we first calculate the covariance of each Pre stimulation and post stimulation trial and return their average then using the averages and whitening matrix and CSP transformation matrix is calculated. This transformation matrix also called mixing matrix is

finally applied to each trial of pre stimulation and post stimulation such that there is an increase in variance of pre MI classes and post MI classes for each subject.

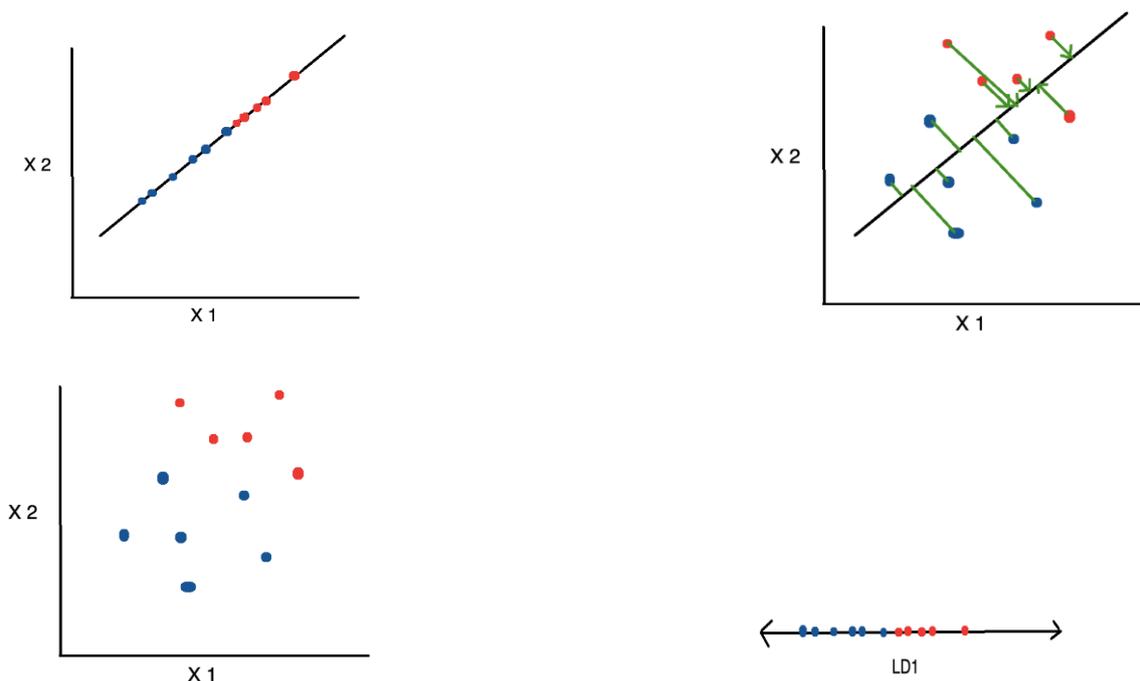
### 3.7 Linear Discriminant Analysis (LDA)

There are classification algorithms limited to only binary classification such as Logistic regression. If the number of classes increases or the classes are not well separated than LDA is preferred technique for classification

This technique is used for dimensionality reduction and classification. It provides classes separation by drawing a decision region between different classes. LDA uses the information from features to create a new axis and projects the data on to the new axis in such a way

- Maximizing the distance between the means of all classes
- Minimizing the variance between each category.

Figure 3.8 shows the different steps used in LDA classification



**Figure 3-8:** LDA classification

## CHAPTER 4: RESULTS AND DISCUSSIONS

### 4.1 Results of Data labeling accuracy of Pre stimulation and Post stimulation data

Different models were trained to analyze the accuracy of output labels. Label 0 represents data before stimulation and label 1 represents data after stimulation. Data of all 10 subjects with all stimulation runs (total of 15 runs for each subject) were first concatenated and then passed to the different models for training and cross validation and results were analyzed on testing data.

**Table 4-1:** Summary of input data for Simple label Classification

<b>Number of observation Rows</b>	<b>1200</b>
<b>Number of features Columns (Predictors):</b>	<b>3279</b>
<b>Response Variable (Labe Column):</b>	<b>3280 nth</b>
<b>Response Classes (0,1):</b>	<b>2</b>
<b>Data Set Size:</b>	<b>31MB</b>
<b>Validation:</b>	<b>Hold out validation with 25% held out</b>

**Table 4-2:** Summary of different models used for classification.

<b>Model</b>	<b>Prediction speed (obs/sec)</b>	<b>Training Time (secs)</b>	<b>Accuracy</b>
<b>Complex Tree</b>	970	7.1527	69.0
<b>Medium Tree</b>	480	5.0076	61.7
<b>Simple Tree</b>	980	4.7398	63
<b>Quadratic Discriminant</b>	1000	5.3959	57.3

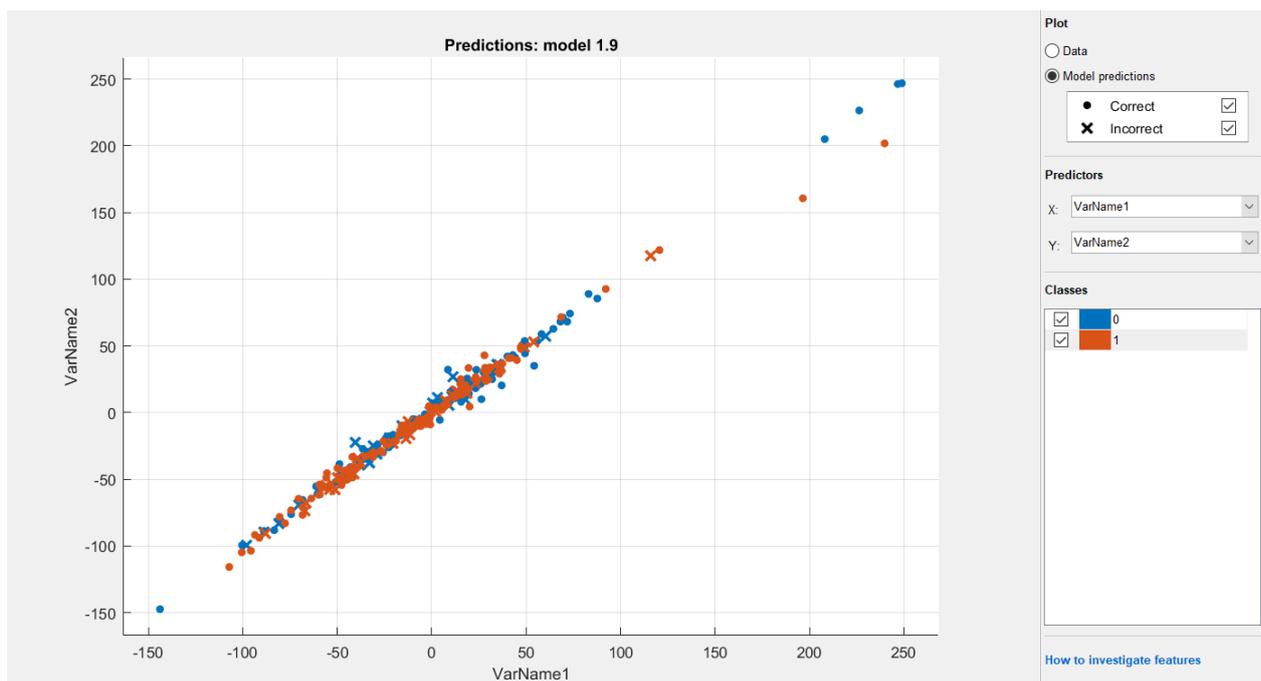
<b>Logistic Regression</b>	1100	6.6296	63.7
<b>Linear SVM</b>	1000	5.5885	63.7
<b>Quadratic SVM</b>	660	6.4391	73.3
<b>Cubic SVM</b>	<b>1100</b>	<b>7.1355</b>	<b>79.7</b>
<b>Fine Gaussian SVM</b>	740	6.3886	62
<b>Medium Gaussian SVM</b>	890	5.6114	55.7
<b>Coarse Gaussian SVM</b>	1000	5.4842	54.7
<b>Fine KNN</b>	<b>1000</b>	<b>5.4466</b>	<b>79.7</b>
<b>Medium KNN</b>	610	5.2785	67.0
<b>Coarse KNN</b>	750	4.8398	57.7
<b>Cosine KNN</b>	890	4.8576	72.7
<b>Cubic KNN</b>	680	4.6173	66.3
<b>Weighted KNN</b>	950	4.8103	76.0
<b>Boosted Trees</b>	660	7.4947	71.3
<b>Bagged Trees</b>	990	6.167	76.3
<b>Subspace Discriminant</b>	480	6.6346	58.3
<b>Subspace KNN</b>	830	6.256	65.7
<b>RUSBoosted Trees</b>	580	6.7401	61.7

Table 4-3 represents 2 models having classification accuracy better than rest of the models used for classification

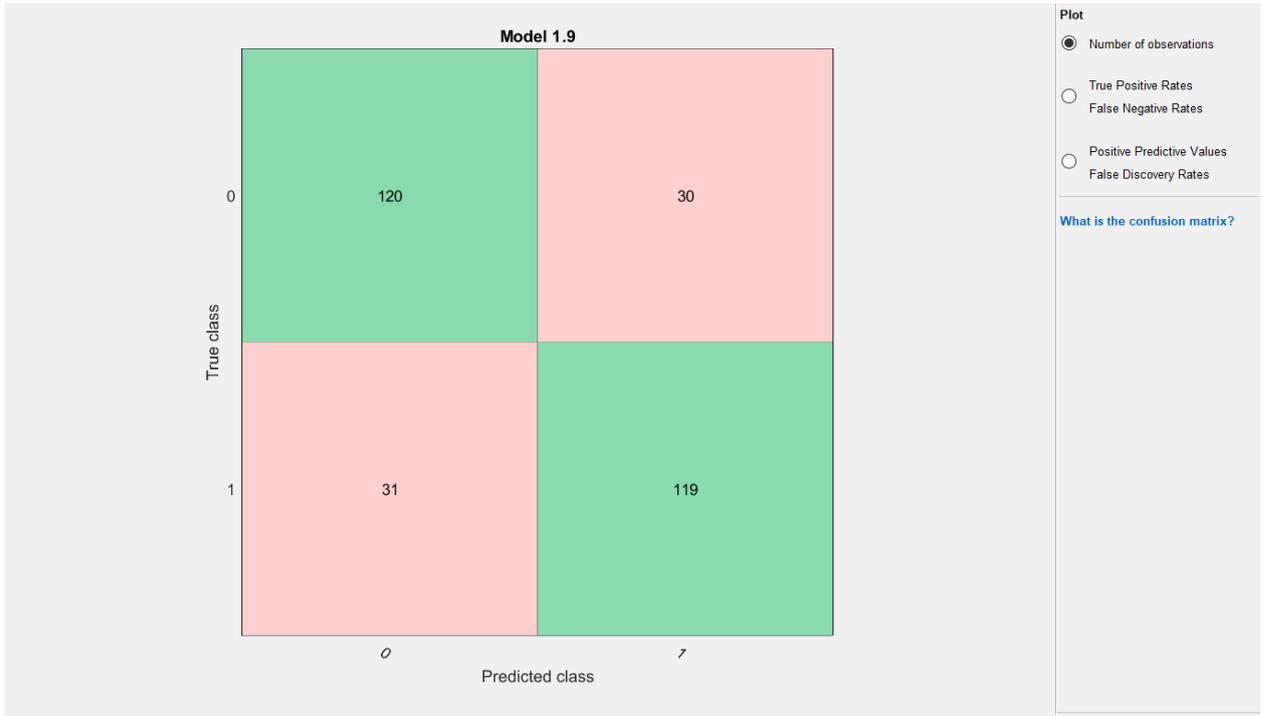
**Table 4-3:** Classification Accuracy of Cubic and Fine KNN models

Model	Classification Accuracy %
Cubic SVM	79.7
Fine KNN	79.7

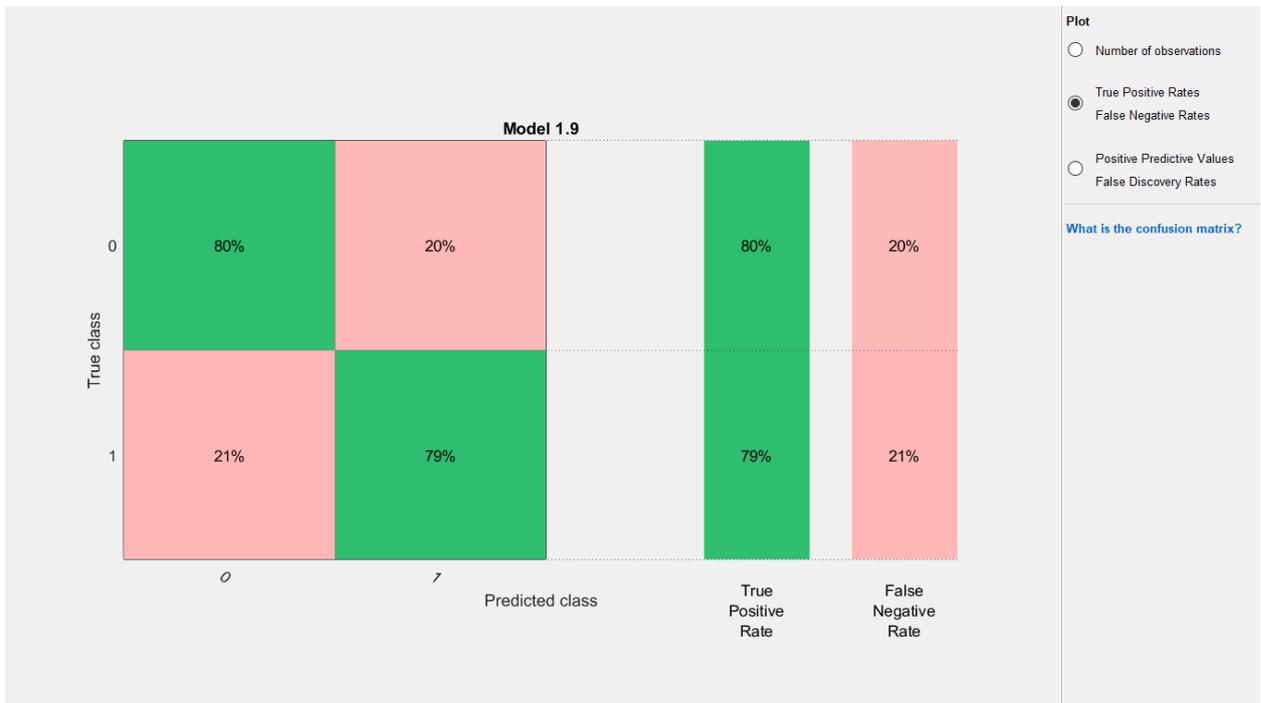
Figure 4.1 shows Scatter plots, figure 4.2 represents confusion matrix, figure 4.3 True Positive and Negative Rate and figure 4.4 shows Positive Predictive Values and False discovery rates for Cubic SVM model used for simple label classification of the data.



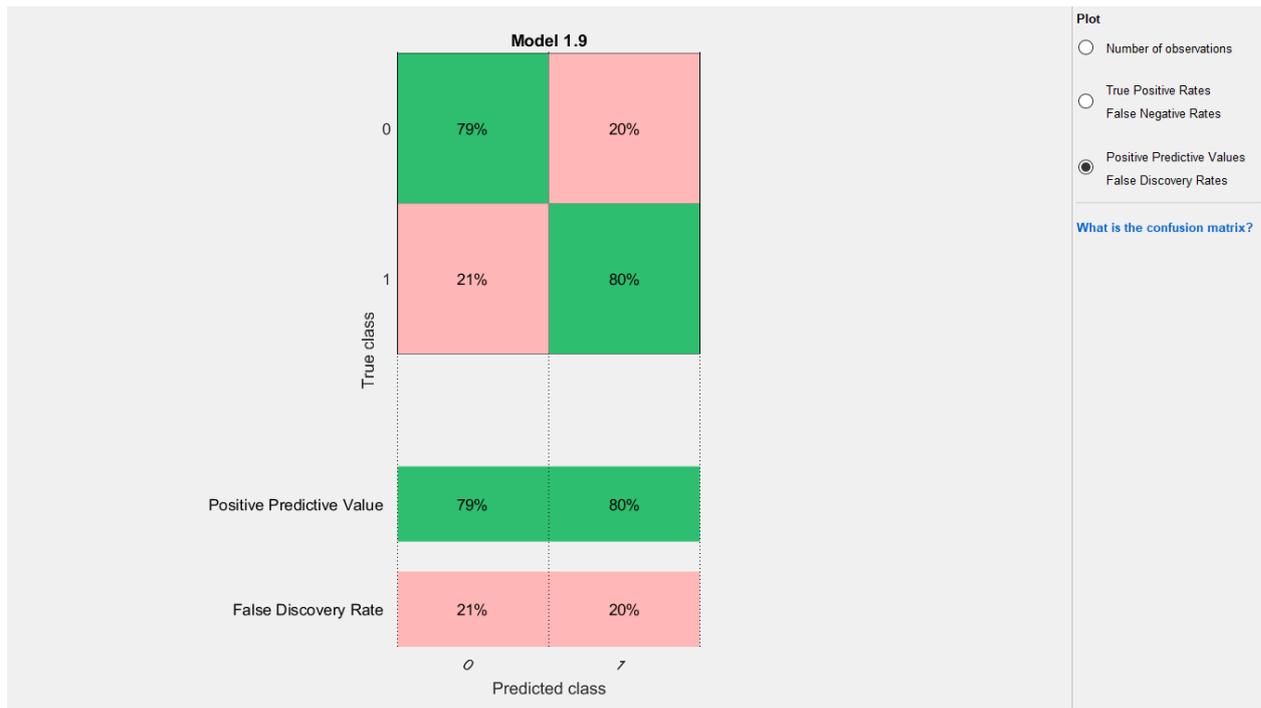
**Figure 4-1:** Scatter Plots of Cubic SVM model



**Figure 4-2:** Confusion Matrix of Cubic SVM model



**Figure 4-3:** True Positive and Negative Rate of Cubic SVM model



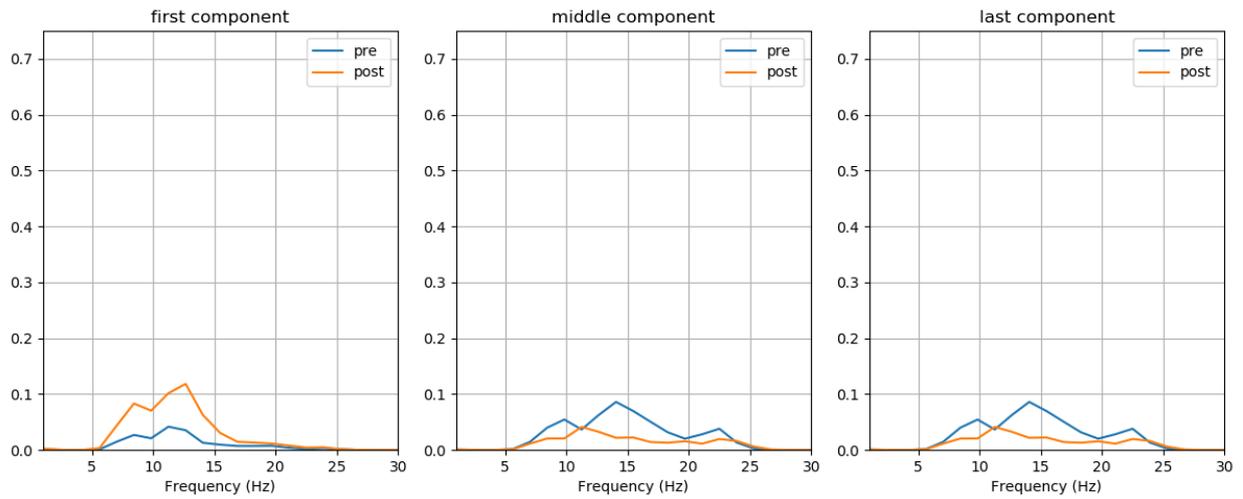
**Figure 4-4:** Positive Predictive Values and False discovery rates of Cubic SVM model

## 4.2 Common Spatial Pattern (CSP)

Common Spatial Pattern (CSP) of pre stimulation and post stimulation runs were plotted for all subjects using python. Following results are drawn from the analysis.

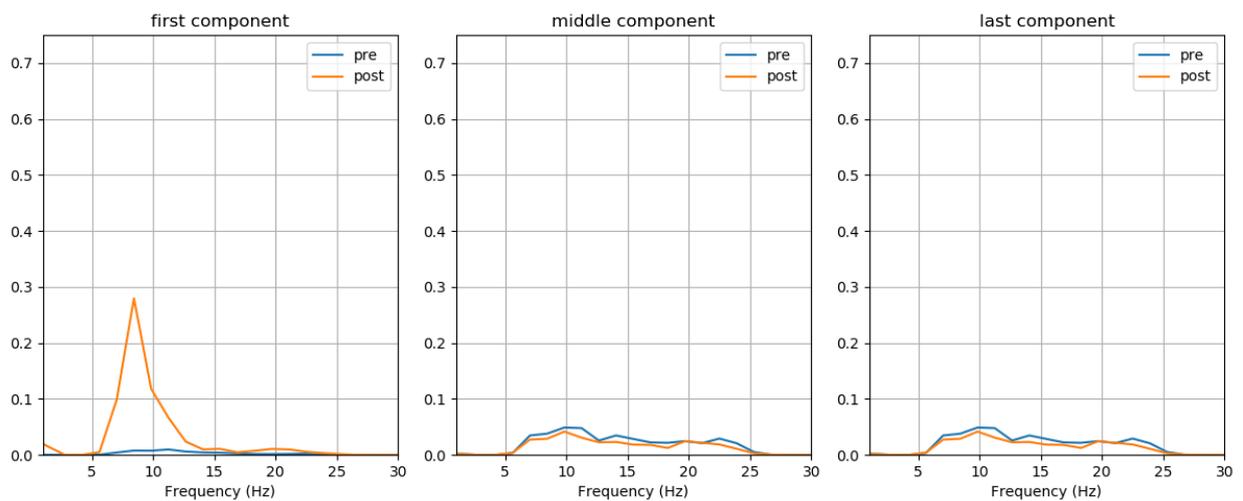
- In CSP we increased the inter variance in pre and post classes of MI epochs
- It is observed by comparison that amplitude of post stimulation MI is higher than pre stimulation MI in both datasets across all subjects.

Figure 4.5 shows CSP of MI for pre stimulation and post stimulation of subject 2 in dataset 1.



**Figure 4-5:** CSP of pre stimulation and post stimulation

Figure 4.6 shows CSP of MI for pre stimulation and post stimulation of subject 2 in dataset 2.



**Figure 4-6:** CSP of pre stimulation and post stimulation

### 4.2.1 Subject 10 CSP

Figure 4.7 shows CSP of MI for pre stimulation and post stimulation of subject 10 in dataset

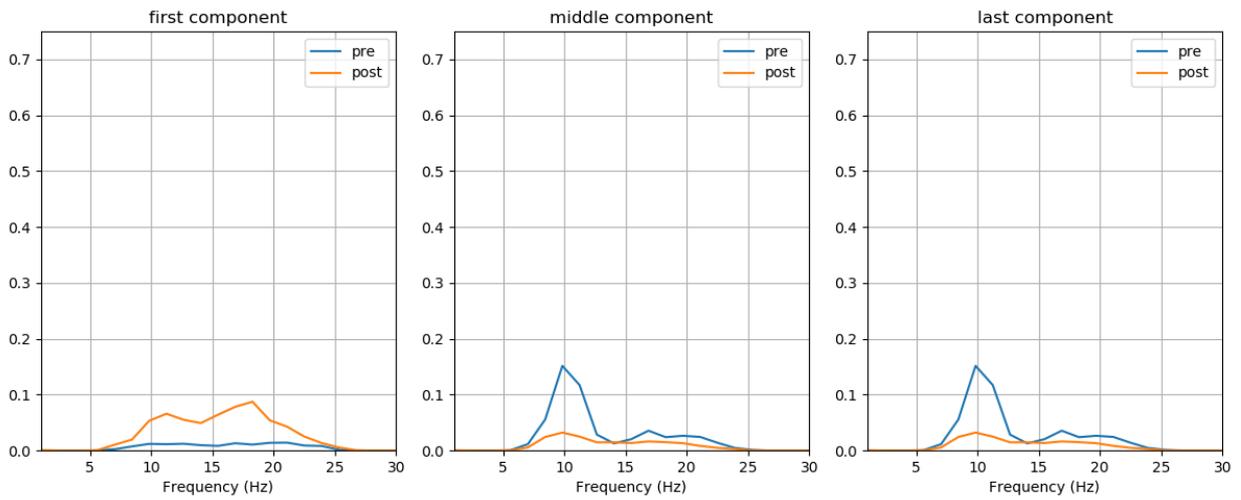


Figure 4-7: CSP of pre stimulation and post stimulation.

Figure 4.8 shows CSP of MI for pre stimulation and post stimulation of subject 9 in dataset 2.

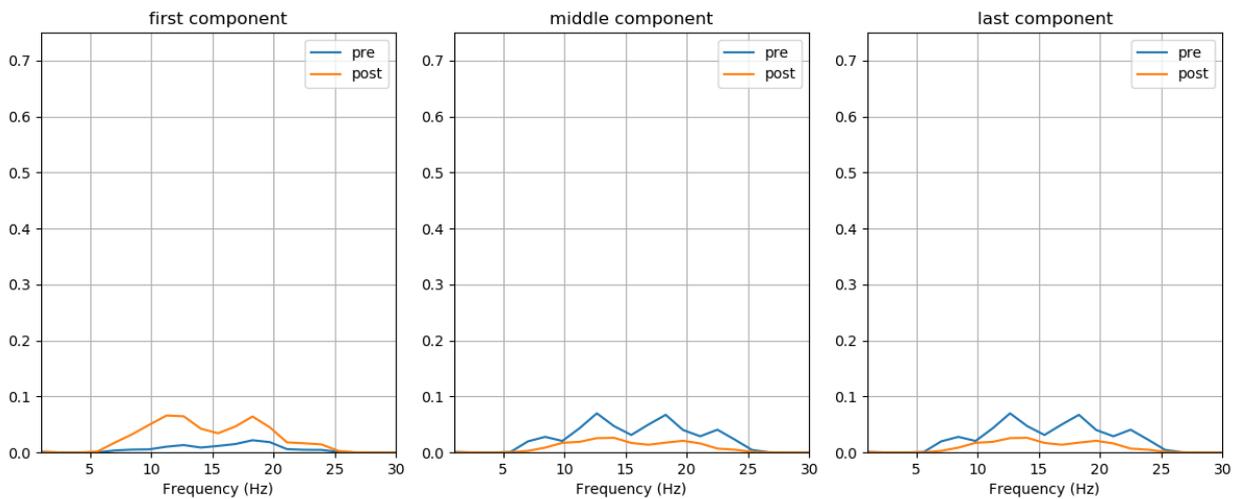
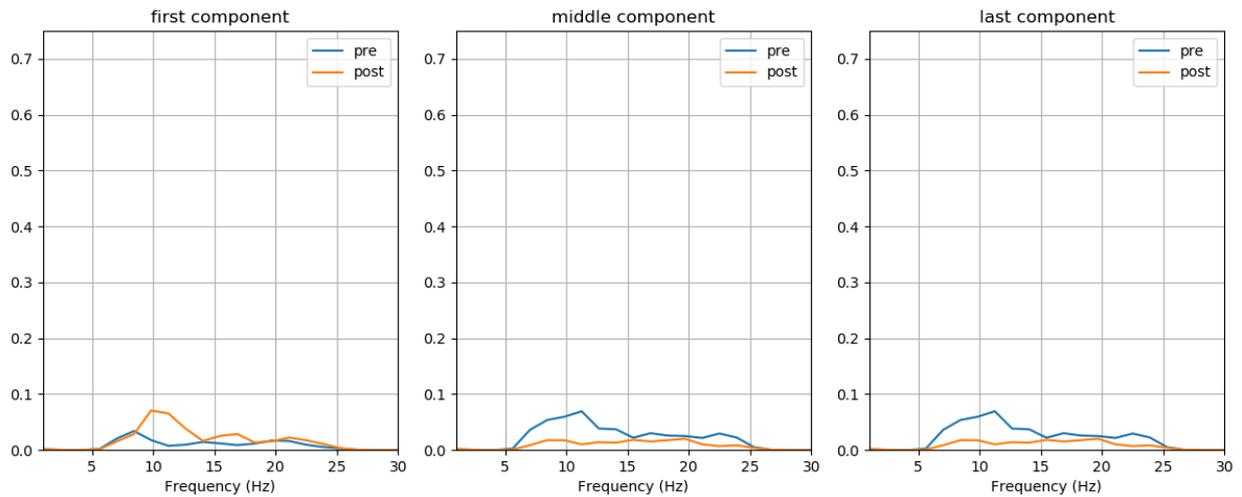


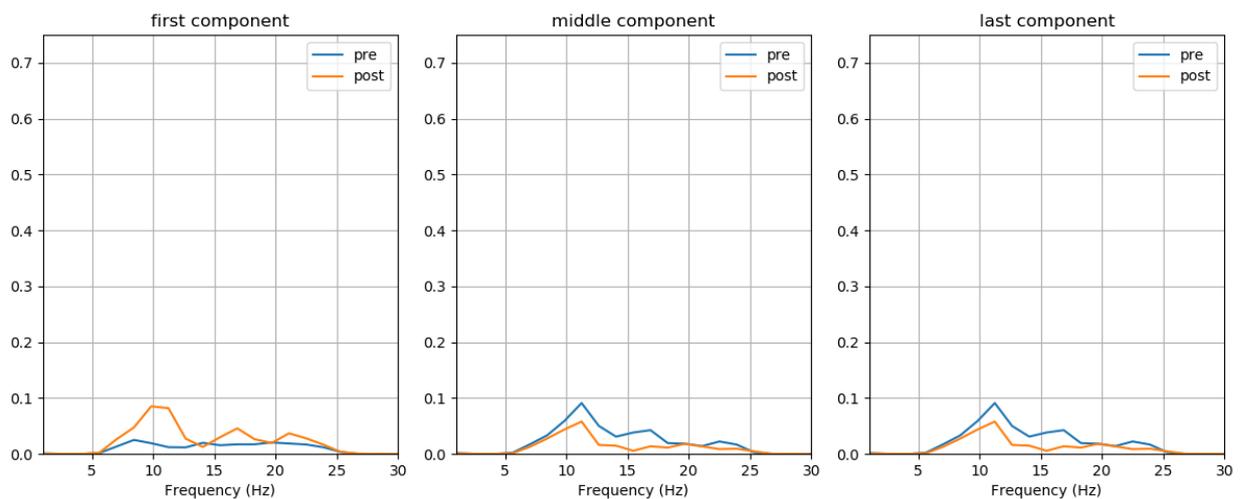
Figure 4-8: CSP of pre stimulation and post stimulation

Figure 4.9 shows CSP of MI for pre stimulation and post stimulation of subject 3 in dataset 1.



**Figure 4-9:** CSP of pre stimulation and post stimulation

Figure 4.10 shows CSP of MI for pre stimulation and post stimulation of subject 3 in dataset 2



**Figure 4-10:** CSP of pre stimulation and post stimulation.

### 4.3 Event Related De synchronization (ERD)

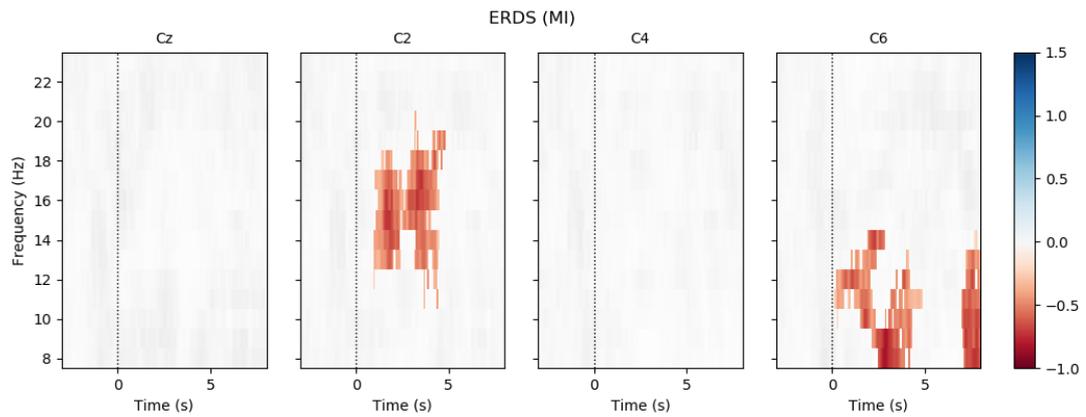
ERDs for all subjects with different channel combination are plotted. The results suggest that the decrease of oscillatory activity related to an event in post stimulation ERDs are better than pre stimulation ERDs as hypothesized for all subjects across all channels however the range of frequency varies, in which amplitude is increased in subjects from 8 Hz to 24 Hz.

The results further suggest contralateral ERDs are reduced during anodal stimulation in post stimulation class than pre stimulation class for seven subjects

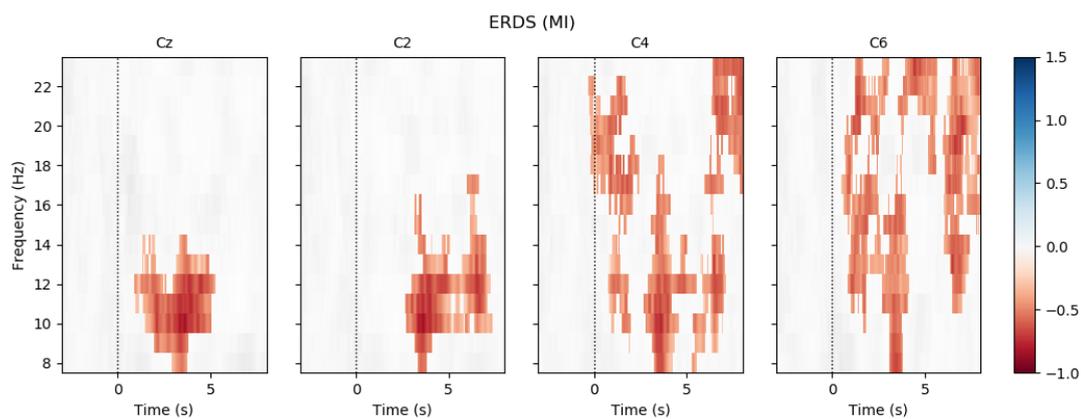
### 4.3.1 ERDs Channels Cz, C2, C4, C6

#### 4.3.1.1 Subject 2

Figure 4.11 and 4.12 shows pre stimulation and post stimulation ERDs for subject 2.



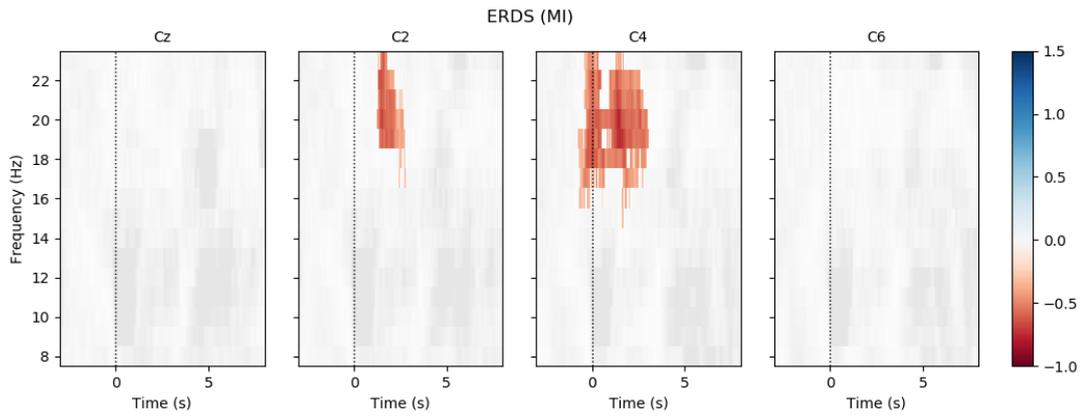
**Figure 4-11:** Pre stimulation ERD for subject 2



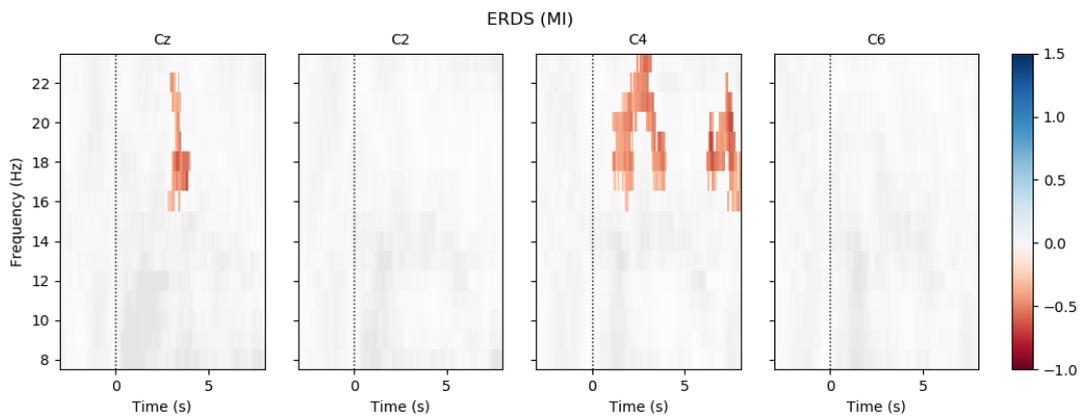
**Figure 4-12:** Post stimulation ERD for subject 2

#### 4.3.1.2 Subject 9

Figure 4.13 and figure 4.14 shows pre stimulation and post stimulation ERDs for subject 9.



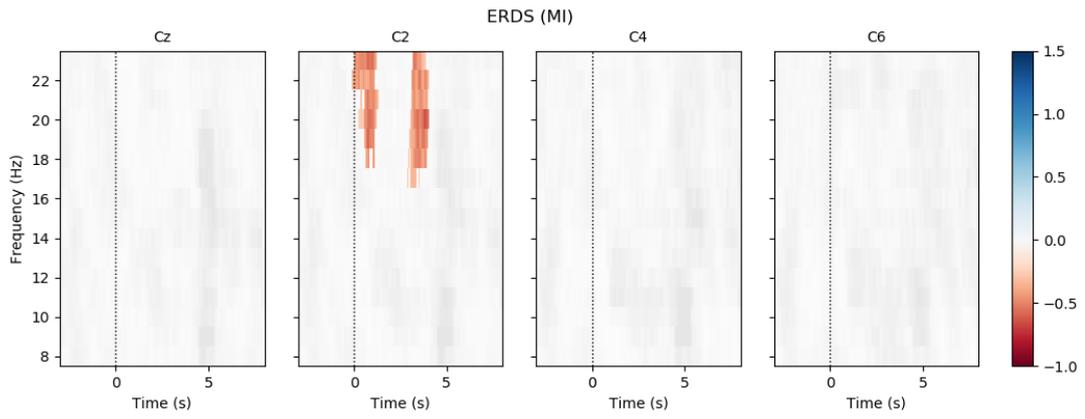
**Figure 4-13:** Pre stimulation ERD for subject 9



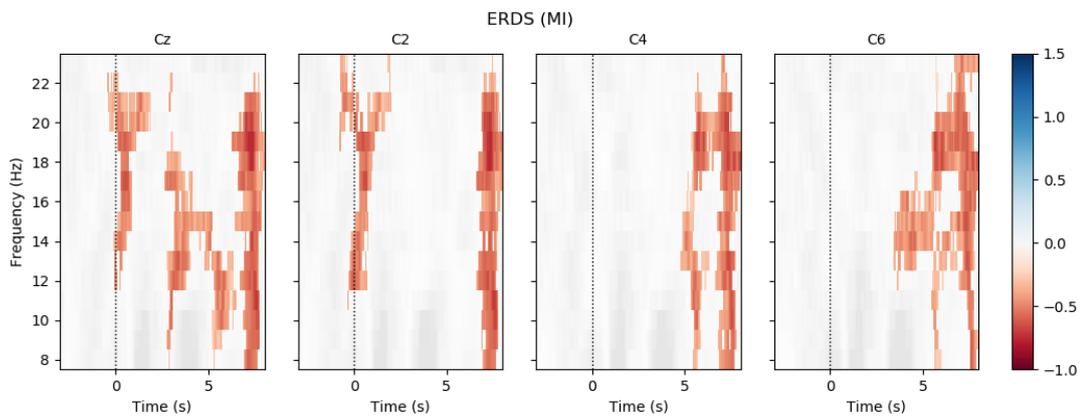
**Figure 4-14:** Post stimulation ERD for subject 9

### 4.3.1.3 Subject 4

Figure 4.15 and figure 4.16 shows pre stimulation and post stimulation ERDs for subject 4.



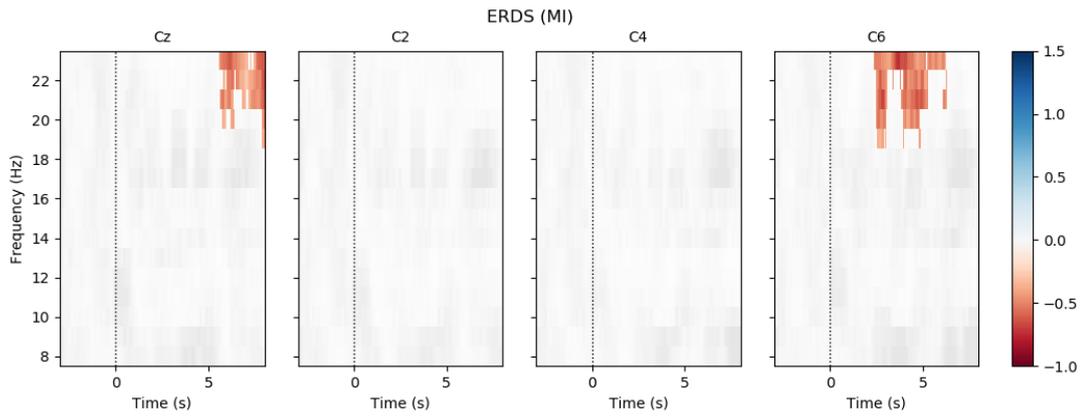
**Figure 4-15:** Pre stimulation ERD for subject 4



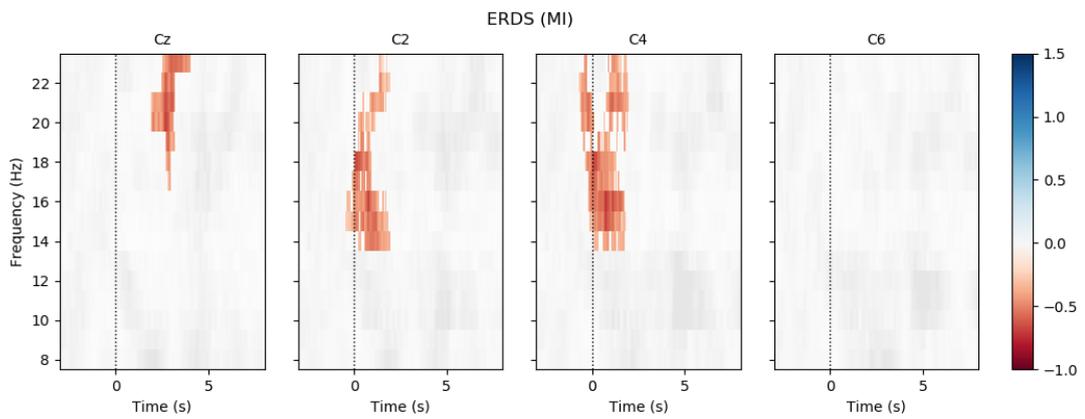
**Figure 4-16:** Post stimulation ERD for subject 4

#### 4.3.1.4 Subject 7

Figure 4.17 and 4.18 shows pre stimulation and post stimulation ERDs for subject 7.



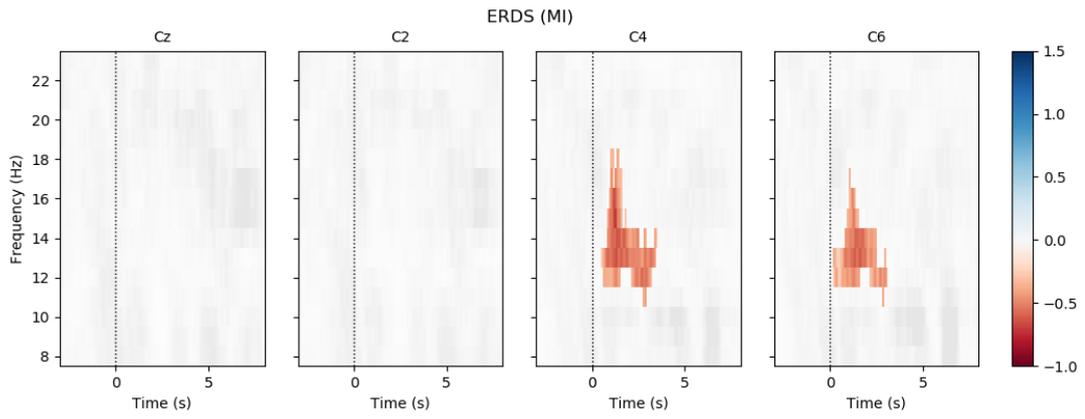
**Figure 4-17:** Pre stimulation ERD for subject 7



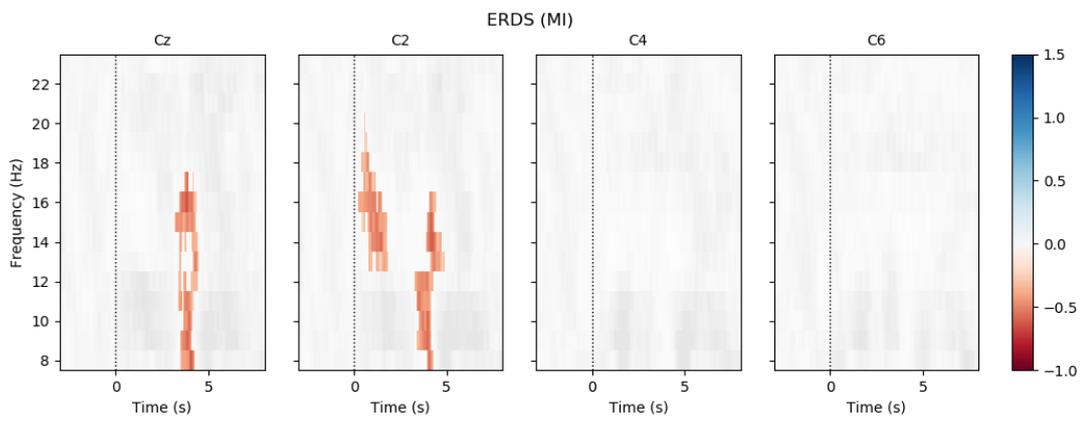
**Figure 4-18:** Post stimulation ERD for subject 7

#### 4.3.1.5 Subject 10

Figure 4.19 and figure 4.20 shows pre stimulation and post stimulation ERDs for subject 10.



**Figure 4-19:** Pre stimulation ERD for subject 10

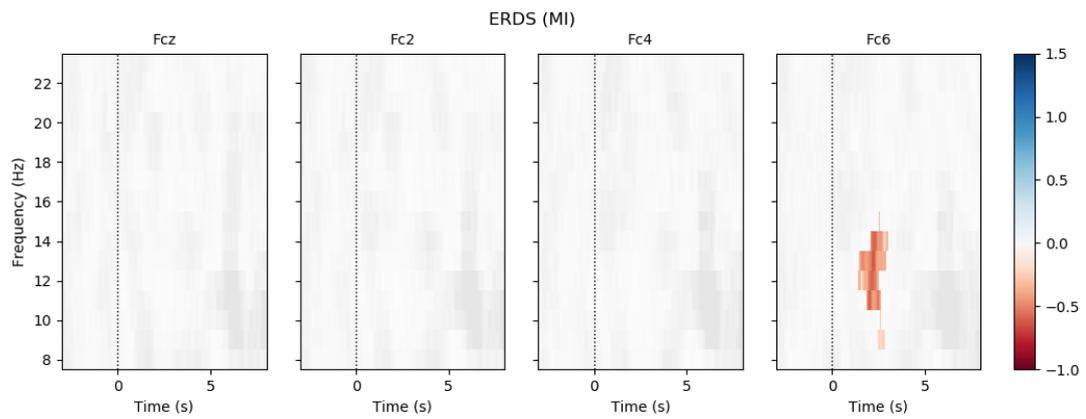


**Figure 4-20:** Post stimulation ERD for subject 10

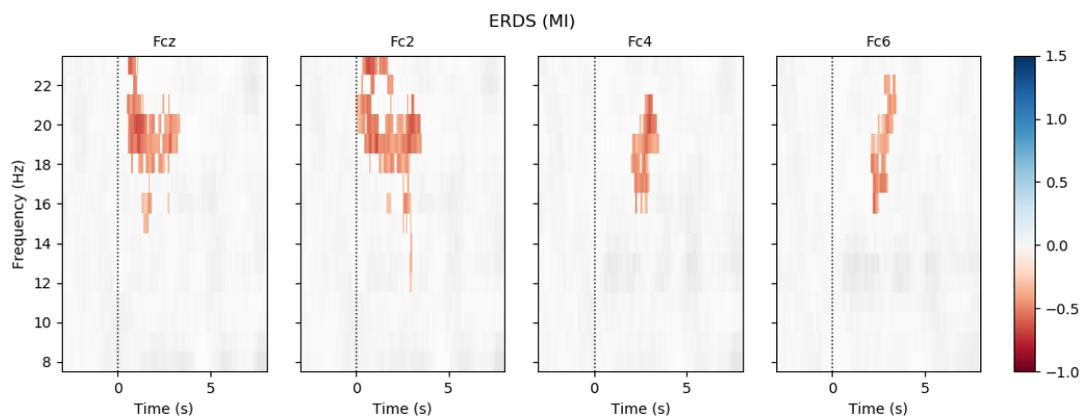
### 4.3.2 Channels FC<sub>Z</sub>, FC2, FC4, FC6

#### 4.3.2.1 Subject 1 ERDs

Figure 4.21 and figure 4.22 shows pre stimulation and post stimulation ERDs for subject 1.



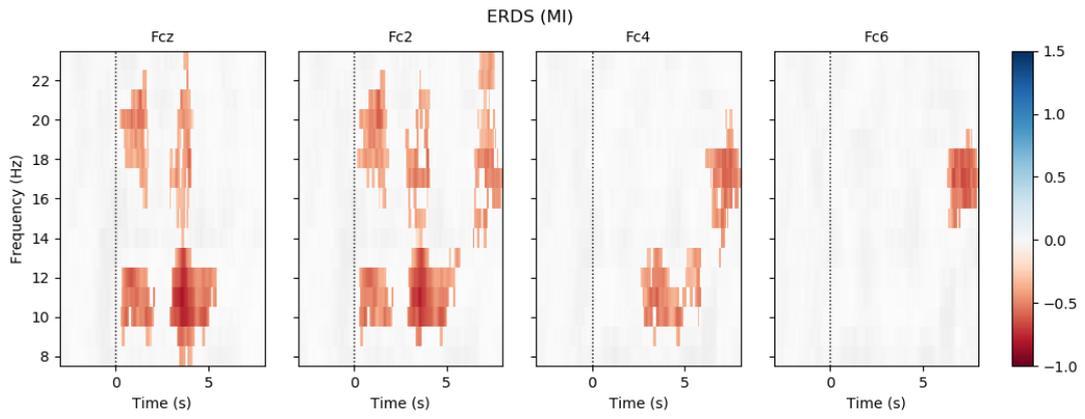
**Figure 4-21:** Pre stimulation ERD for subject 1



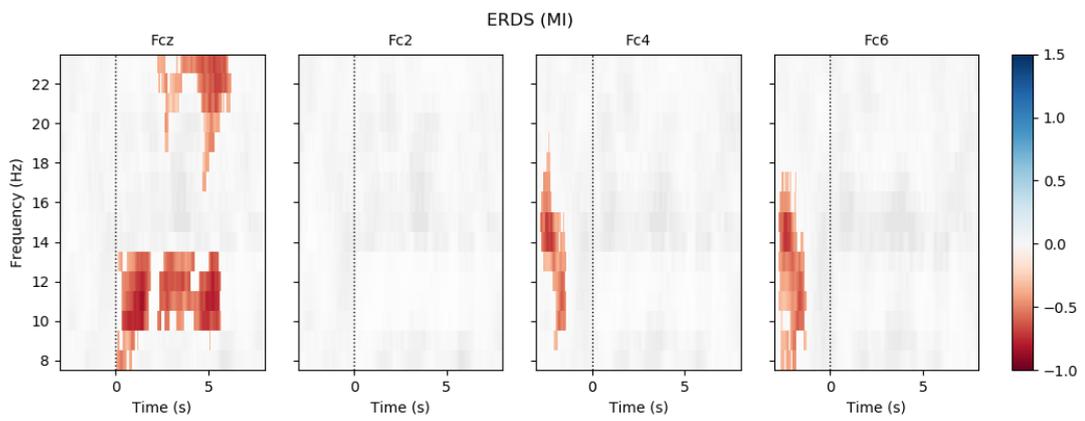
**Figure 4-22:** Post stimulation ERD for subject 1

#### 4.3.2.2 Subject 3

Figure 4.213and figure 4.24 shows pre stimulation and post stimulation ERDs for subject 3.



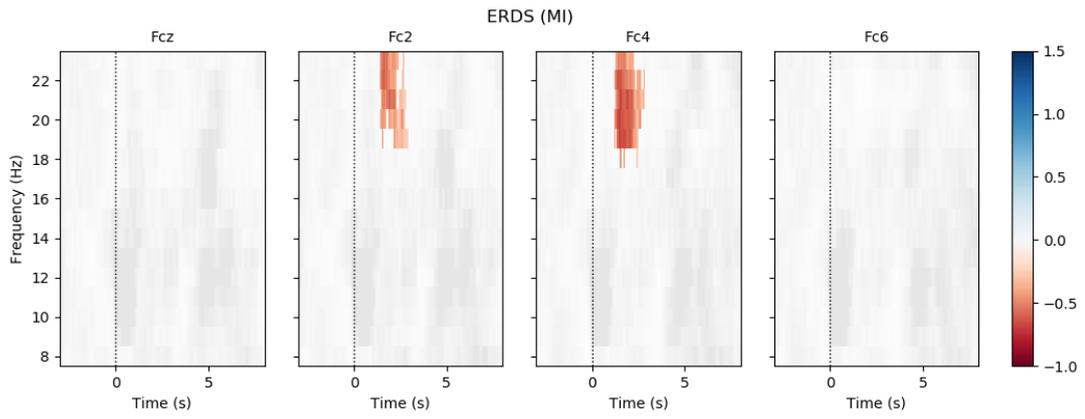
**Figure 4-23:** Pre stimulation ERD for subject 3.



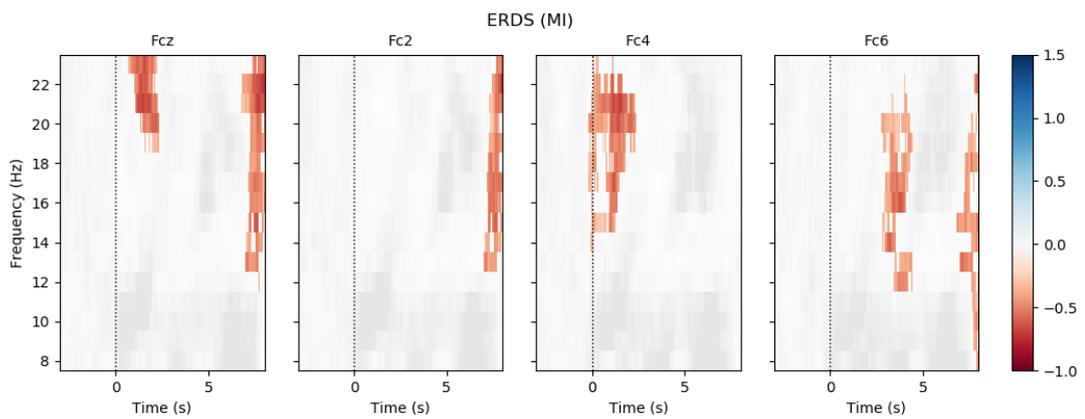
**Figure 4-24:** Post stimulation ERD for subject 3.

### 4.3.2.3 Subject 9

Figure 4.25 and figure 4.26 shows pre stimulation and post stimulation ERDs for subject 9.



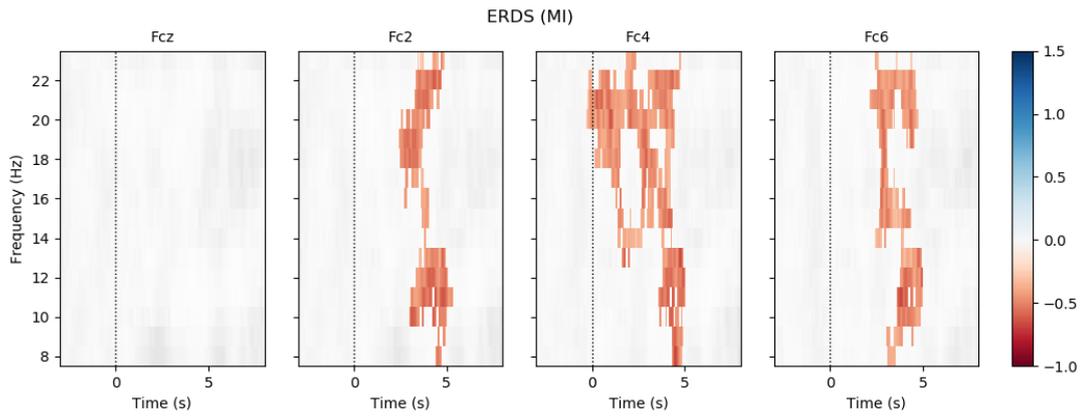
**Figure 4-25:** Pre stimulation ERD for subject 9



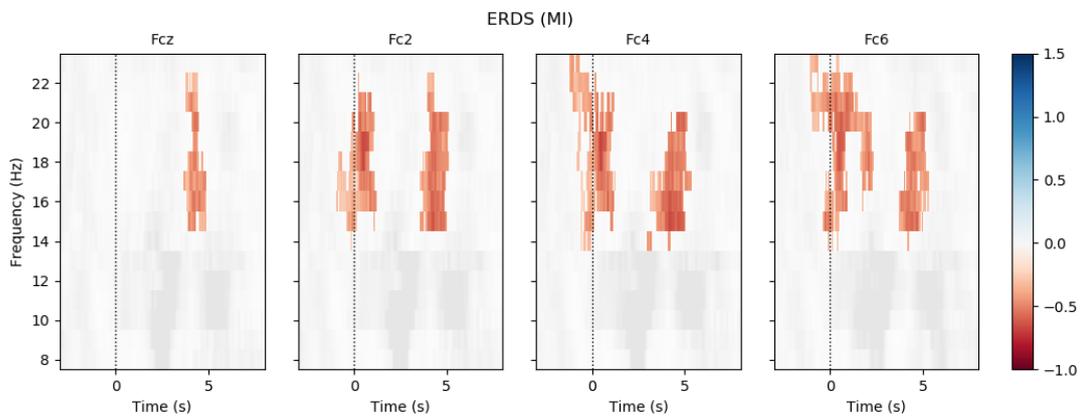
**Figure 4-26:** Post stimulation ERD for subject 9

#### 4.3.2.4 Subject 7

Figure 4.27 and figure 4.28 shows pre stimulation and post stimulation ERDs for subject 7.



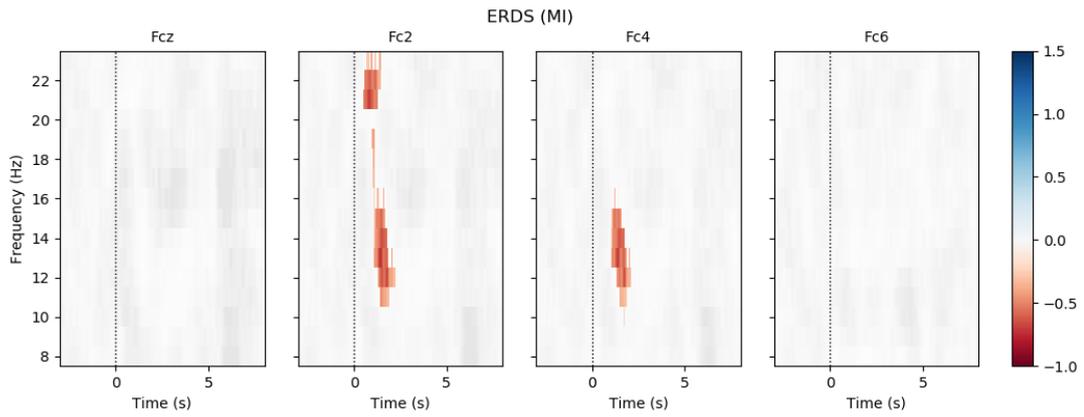
**Figure 4-27:** Pre stimulation ERD for subject 7



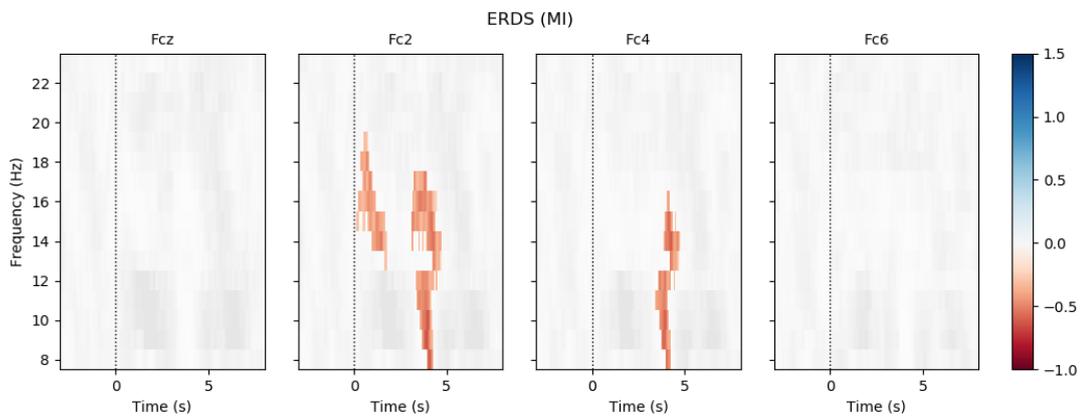
**Figure 4-28:** Post stimulation ERD for subject 7

#### 4.3.2.5 Subject 10

Figure 4.29 and figure 4.30 shows pre stimulation and post stimulation ERDs for subject 10.



**Figure 4-29:** Pre stimulation ERD for subject 10



**Figure 4-30:** Post stimulation ERD for subject 10

#### 4.4 Linear Discriminant Analysis (LDA) Classification

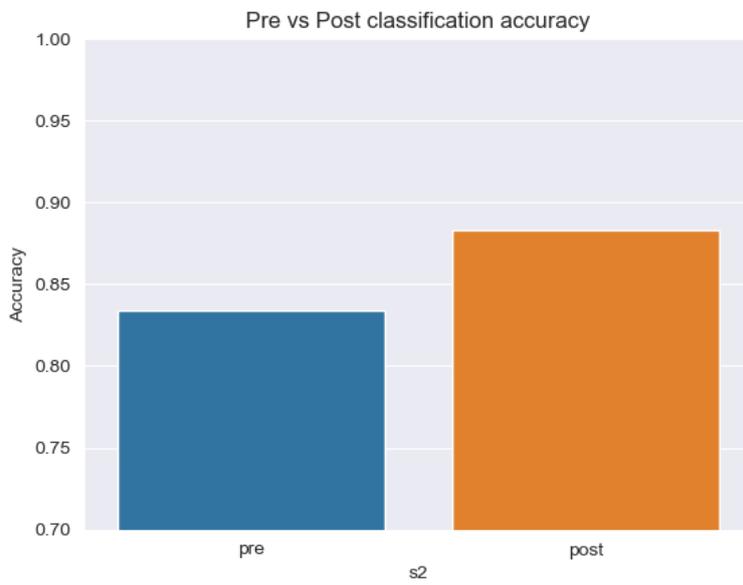
The results are further validated by Linear Discriminant Analysis (LDA) a machine learning classification model on CSP and ERDS of pre and post stimulation data. Table 4-4 shows LDA motor imagery classification accuracies on CSP for dataset 2

**Table 4-4:** LDA Motor Imagery Detection Accuracy dataset 2

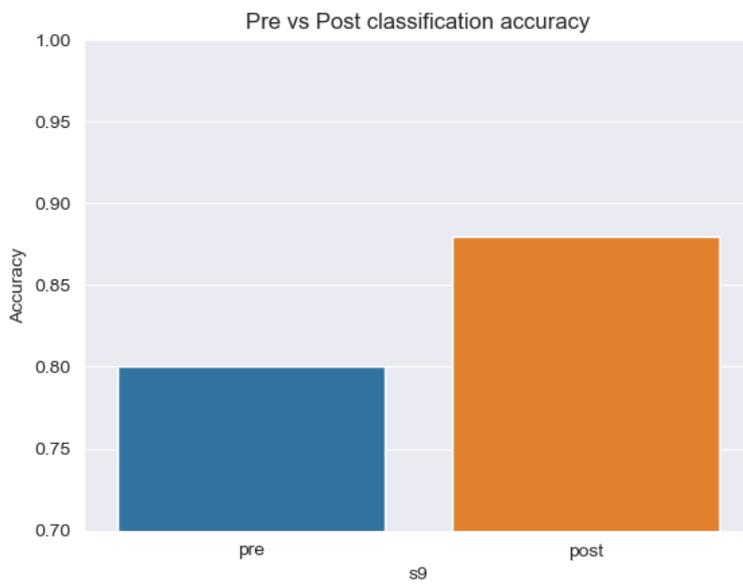
<b>Subject</b>	<b>Pre Stimulation MI Accuracy</b>	<b>Post Stimulation MI Accuracy</b>
<b>S1+</b>	0.825	0.901
<b>S2+</b>	0.834	0.883
<b>S3</b>	0.892	0.88
<b>S4+</b>	0.81	0.883
<b>S5+</b>	0.861	0.893
<b>S6</b>	0.866	0.855
<b>S7+</b>	0.88	0.943
<b>S8</b>	0.874	0.830
<b>S9+</b>	0.803	0.882
<b>S10+</b>	0.890	0.961

It is observed in dataset 2 that classification of Motor imagery period (MIP) and rest period (RP) after the stimulation therapy is higher in seven subjects out of ten compare to the detection of (MIP) and rest period (RP) of pre stimulation data.

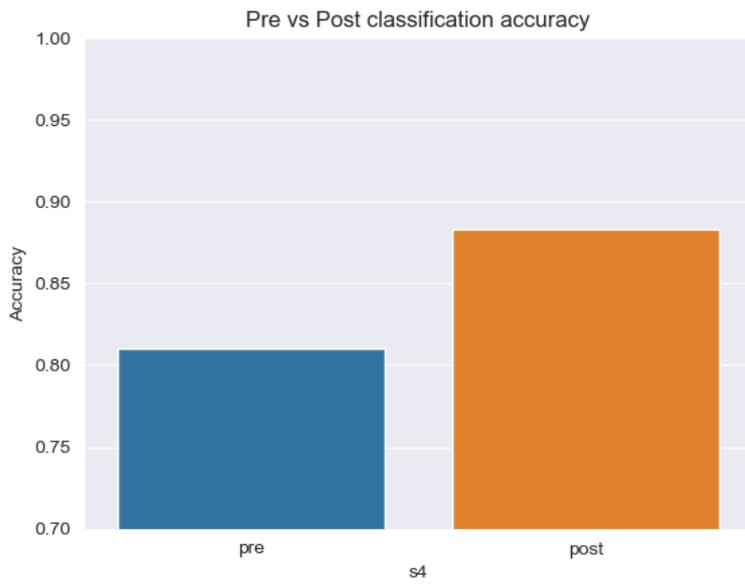
Figure 4.31 to 4.34 compares Pre and post MI classification accuracies of subjects on dataset



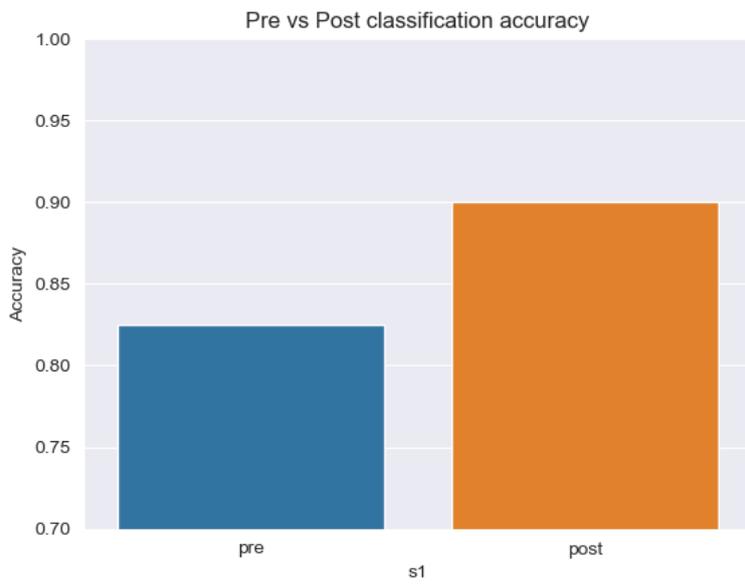
**Figure 4-31:** Subject 2 MI Classification Accuracy



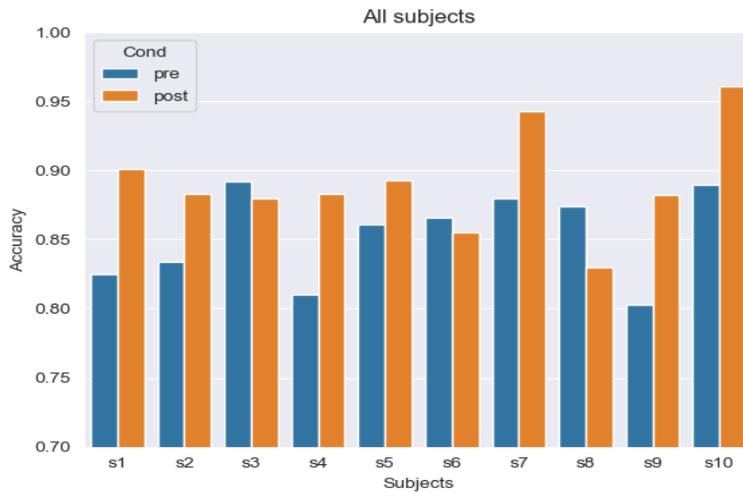
**Figure 4-32:** Subject 1 MI Classification Accuracy



**Figure 4-33:** Subject 9 MI Classification Accuracy



**Figure 4-34:** Subject 4 MI Classification Accuracy



**Figure 4-35:** MI classification comparison on CSP

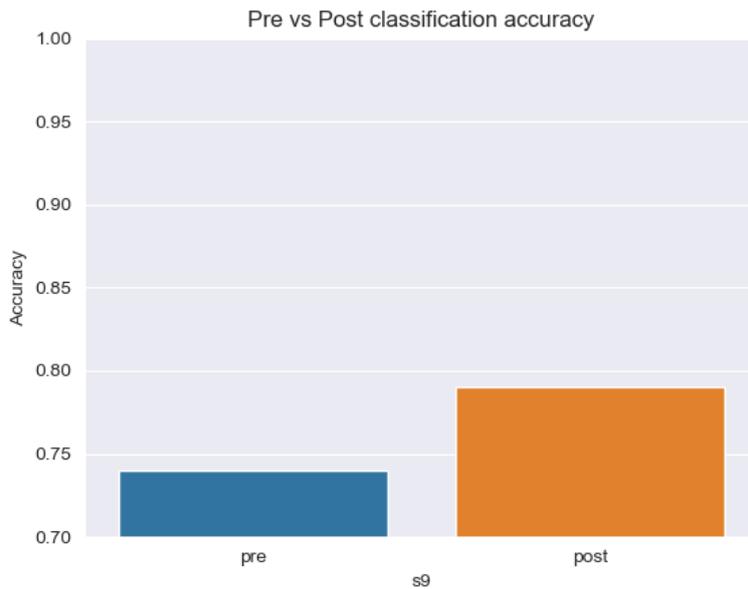
Table 4-5 shows LDA motor imagery classification accuracies on CSP for dataset 1

**Table 4-5:** LDA Motor Imagery Detection Accuracy dataset 1

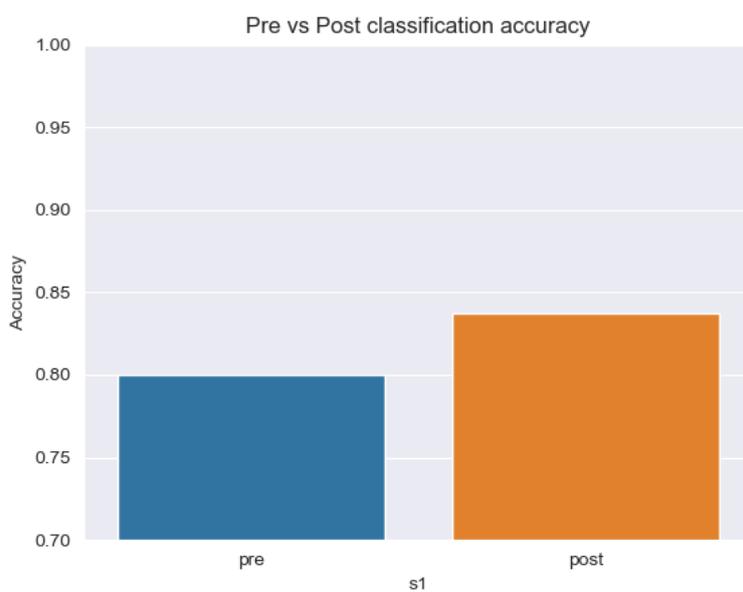
<b>Subject</b>	<b>Pre Stimulation MI Accuracy</b>	<b>Post Stimulation MI Accuracy</b>
<b>S1+</b>	0.801	0.837
<b>S2+</b>	0.766	0.839
<b>S3</b>	0.840	0.821
<b>S4</b>	0.800	0.782
<b>S5+</b>	0.791	0.842
<b>S6</b>	0.805	0.733
<b>S7+</b>	0.82	0.87
<b>S8</b>	0.87	0.84
<b>S9+</b>	0.74	0.79
<b>S10+</b>	0.83	0.85

Similarly for dataset 1 classification of Motor imagery period (MIP) and rest period (RP) after the stimulation therapy is higher in six subjects out of ten compared to the detection of (MIP) and rest period (RP) of pre stimulation data.

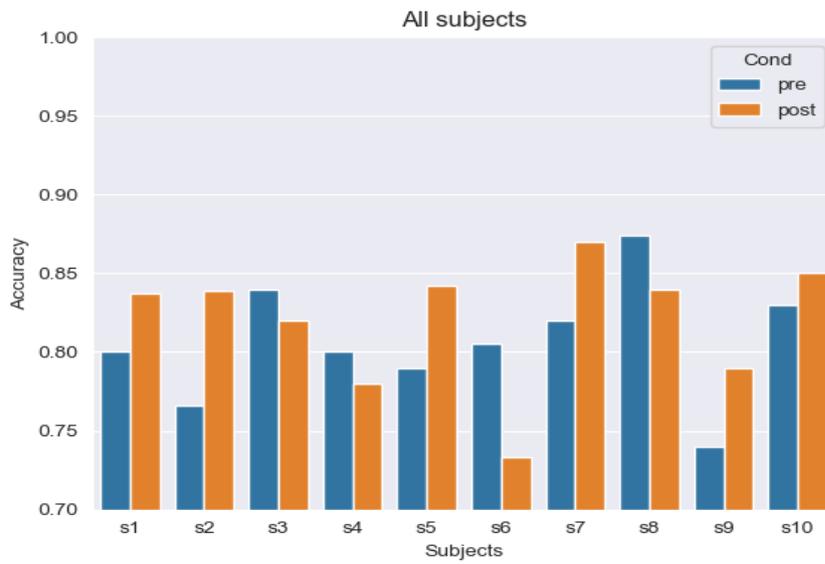
Figure 4.36 to 4.38 compares Pre and post MI classification accuracies of subjects on dataset.



**Figure 4-36:** Subject 1 MI Classification Accuracy



**Figure 4-37:** Subject 9 MI Classification Accuracy



**Figure 4-38:** MI classification comparison on CSP

## **CHAPTER 5: CONCLUSIN AND FUTURE WORK**

### **5.1 Conclusion**

This study proposes a method of classifying Motor Imagery Period (MIP) and Rest Period (RP) in the Electroencephalogram (EEG) recorded signals based on Common Spatial Patterns (CSP) and Event Related De synchronization ERD. Further Linear Discriminant Analysis (LDA) classification is applied to these patterns on two datasets. Dataset 2 is online available dataset and dataset 1 is recorded in the lab for the study.

Following are the conclusion of this study:

- The findings have shown that detection of Motor imagery in comparison with rest period in the whole EEG recorded signal improves after the stimulation therapy, which is indication of cognitive state improvement
- ERD and CSP features are used and validated by LDA classifier.
- Higher Post MI classification accuracy is estimated in 7 and 5 subject using Dataset 2 and 1 respectively out of 10 subjects
- Noninvasive therapy such as tDCS has shown affective changes in the cognitive improvement and is estimated by dataset I and II.

### **5.2 Future Work**

The finding of this study can be extended to analyze similar results using different techniques other than CSP, ERD and LDA. The recorded dataset can further be analyzed using statistical analysis techniques such as t test. The recording session for each subject and number of subjects can be increased with different Electrodes montages.

The study can then be further extended by varying the intensity of direct current stimulation and electrodes placement to find the efficacy of the noninvasive therapy for different settings and the results be compared.

## APPENDIX 1 ERD

```
import mne

import numpy as np

import matplotlib as mp

import pandas as pd

from mne import Epochs, pick_types, events_from_annotations

from mne.channels import read_layout

from mne.preprocessing import ICA

import matplotlib.pyplot as plt

from mne.time_frequency import tfr_multitaper

from mne.stats import permutation_cluster_1samp_test as pcluster_test

from mne.viz.utils import center_cmap

from mne.io import concatenate_raws

import os

import glob

frame=pd.DataFrame()

dataarray=np.zeros([10,14,15000])# i changed this value from 15000

df=pd.DataFrame()

basepath='S05/day2'

iter=0

for trail in os.listdir(os.path.join(basepath)):

    for file in os.listdir(os.path.join(basepath,trail)):
```

```

if(trail[:-2]=='neurofeedback' and (int(trail[-2:])<=5 or int(trail[-2:])>=11)):

    df=pd.read_csv(os.path.join(basepath,trail,file),sep='\t',header=None)

    df=df.T

    print(basepath,trail,file)

    print(df.shape)

    data=df.values;

    dataarray[iter,:,:]=data[:,0:15000] # i changed this value 15000

    iter+=1

dataarray.shape

dataarray[:,:13,:]=dataarray[:,:13,:]/1000000

data=dataarray

[ntrail,nchann,nsamp]=data.shape

sampling_freq = 128 # in Hertz

ch_names =['F4', 'F8', 'C4', 'T4', 'Cz', 'C2', 'T6', 'C6', 'CPz', 'CP2', 'CP4', 'CP6','A1','STI 001']

#ch_names = ['FCz', 'FC2', 'FC4', 'FC6', 'Cz', 'C2', 'C4', 'C6', 'CPz', 'CP2', 'CP4',
'CP6','A1','STI 001']

ch_types = ['eeg', 'eeg', 'stim']

info = mne.create_info(ch_names=ch_names, sfreq=sampling_freq, ch_types=ch_types)

raw= mne.io.RawArray(data[6,:,:], info)***change this value one by one, Save Figure 3
ONLY for Pre MI 0 to 4, for Post MI,5 t0 9

raw.filter(l_freq=3, h_freq=40) #filter the fraw data from 8 to 22 hertz

raw.set_montage('standard_1020')

raw.plot();

```

```

#raw = concatenate_raws([mne.io.RawArray(data[f,:,:], info) for f in range(0,ntrail)])
# raw=mne.io.concatenate_raws([raw, mne.io.RawArray(data[1,:,:], info)])

# ica = mne.preprocessing.ICA(n_components=13, random_state=97, max_iter=800)
# ica.fit(raw)
# ica.plot_sources(raw, show_scrollbars=True);
# ica.plot_components(title='compnents',cmap='jet')

# ica.plot_properties(raw, picks=[0,1,2,3,4,5,6,7,8, 9,10,11])

# print(raw.info)
# #raw.filter(l_freq=12, h_freq=22)
# # raw.set

events = mne.find_events(raw, 'STI 001')
print(events)

raw.info['bads'] += [ 'A1' ]
picks = mne.pick_types(raw.info, eeg=True, stim=False,exclude='bads')
picks = mne.pick_channels(raw.info["ch_names"], ['F4', 'F8','C4', 'T4',])
tmin, tmax = -3, 8 # define epochs around events (in s)
    event_ids = {'ready': 1, 'MI': 2, 'postrest': 3} # map event IDs to tasks
reject= dict( eeg=150e-6 ) # 150 µV
epochs = mne.Epochs(raw, events, event_ids, tmin - 0.5, tmax + 0.5,
    picks=picks,baseline=None, preload=True)

```

```

# ica = mne.preprocessing.ICA(n_components=12, random_state=97, max_iter=800)

# ica.fit(epochs)

# ica.plot_sources(epochs, show_scrollbars=True);

# ica.plot_components(title='components', cmap='jet')

freqs = np.arange(8, 24, 1) # frequencies from 2-35Hz
n_cycles = freqs # use constant t/f resolution
vmin, vmax = -1, 1.5 # set min and max ERDS values in plot
baseline = [-3, 0] # baseline interval (in s)
cmap = center_cmap(plt.cm.RdBu, vmin, vmax) # zero maps to white
kwargs = dict(n_permutations=100, step_down_p=0.05, seed=1,
              buffer_size=None, out_type='mask') # for cluster test

# Run TF decomposition overall epochs
tfr = tfr_multitaper(epochs, freqs=freqs, n_cycles=n_cycles,
                    use_fft=True, return_itc=False, average=False,
                    decim=2)
tfr.crop(tmin, tmax)
tfr.apply_baseline(baseline, mode="percent")
for event in event_ids:
    # select desired epochs for visualization
    tfr_ev = tfr[event]

```

```
fig, axes = plt.subplots(1,5,figsize=(12, 4), gridspec_kw={"width_ratios": [10, 10, 10,10,10]})
```

```
for ch, ax in enumerate(axes[:-1]): # for each channel

    # positive clusters
    _, c1, p1, _ = pcluster_test(tfr_ev.data[:, ch, ...], tail=1, **kwargs)

    # negative clusters
    _, c2, p2, _ = pcluster_test(tfr_ev.data[:, ch, ...], tail=-1,
                                **kwargs)

    # note that we keep clusters with  $p \leq 0.05$  from the combined clusters
    # of two independent tests; in this example, we do not correct for
    # these two comparisons
    c = np.stack(c1 + c2, axis=2) # combined clusters
    p = np.concatenate((p1, p2)) # combined p-values
    mask = c[..., p <= 0.05].any(axis=-1)

    # plot TFR (ERDS map with masking)
    tfr_ev.average().plot([ch], vmin=vmin, vmax=vmax, cmap=(cmap, False),
                          axes=ax, colorbar=False, show=False, mask=mask,
                          mask_style="mask")

    ax.set_title(epochs.ch_names[ch], fontsize=10)
    ax.axvline(0, linewidth=1, color="black", linestyle=":") # event
    if ch != 0:
        ax.set_ylabel("")
```

```

        ax.set_yticklabels("")
fig.colorbar(axes[0].images[-1], cax=axes[-1])
fig.suptitle("ERDS ({}).format(event))
fig.show()

```

## APPENDIX II CSP

```

import mne

import numpy as np

#import matplotlib as mp

import pandas as pd

from mne import Epochs, pick_types, events_from_annotations

from mne.channels import read_layout

from mne.preprocessing import ICA

from mne.io import concatenate_raws

import matplotlib.pyplot as plt

from matplotlib import mlab

import os

import glob

frame=pd.DataFrame()

dataarray=np.zeros([10,14,15000])# i changed this value from 15000

df=pd.DataFrame()

basepath='S01/day2'

iter=0

for trail in os.listdir(os.path.join(basepath)):

    for file in os.listdir(os.path.join(basepath,trail)):

```

```

if(trail[:-2]=='neurofeedback' and (int(trail[-2:])<=5 or int(trail[-2:])>=11)):
    df=pd.read_csv(os.path.join(basepath,trail,file),sep='\t',header=None)
    df=df.T
    print(basepath,trail,file)
    print(df.shape)
    data=df.values;
    dataarray[iter,,:]=data[:,0:15000] # i changed this value 15000
    iter+=1

dataarray.shape

eegdata=dataarray

cols=['Fcz', 'Fc2', 'Fc4', 'Fc6', 'Cz', 'C2', 'C4', 'C6', 'Cpz', 'Cp2', 'Cp4', 'Cp6','Cz','stim']

eegdata[:,13,:]=eegdata[:,13,:]/1000000

#df=pd.DataFrame(data=data.T,columns=cols)

#print(eegdata[1,2,:20])

#datastack=np.reshape(data,(10*14,18000))

sampling_freq = 128 # in Hertz

Fs=sampling_freq

[ntrail,nchann,nsamp]=eegdata.shape

%matplotlib

sampling_freq = 128 # in Hertz

ch_names = ['FCz', 'FC2', 'FC4', 'FC6', 'Cz', 'C2', 'C4', 'C6', 'CPz', 'CP2', 'CP4', 'CP6','A1','STI
001']

ch_types = ['eeg', 'eeg', 'stim']

```

```

info = mne.create_info(ch_names=ch_names, sfreq=sampling_freq, ch_types=ch_types);
#for trail in range(0,ntrail):

raw= mne.io.RawArray(eegdata[0,:,:], info) # reading neurofeddback 2 which correspond to
prestim 3 n3

raw.set_montage("standard_1020")

#raw.set_eeg_reference(['Cz'])

raw.filter(l_freq=8, h_freq=22) #filter the fraw data from 8 to 22 hertz

#raw.plot_psd(fmax=50); #plot psd of

start, stop = raw.time_as_index([100, 115]) # 100 s to 115 s data segment

data, times = raw[:, start:stop]

#plt.plot(raw.get_data()[0])

#data, times = raw[2:20:3, start:stop] # access underlying data

#plotraw.get_data()[0]

raw.plot(); #plot raw data pre

reject= dict( eeg=150e-6 ) # 150 µV

event_id = {'ready': 1, 'MI': 2, 'postrest': 3}

events = mne.find_events(raw, 'STI 001');

#fig = mne.viz.plot_events(events, event_id=id,
sfreq=raw.info['sfreq'],first_samp=raw.first_samp)

```

```

raw.info['bads'] += [ 'A1']

picks = mne.pick_types(raw.info, eeg=True, stim=False, exclude='bads')

epochs = mne.Epochs(raw, events, event_id, tmin=-0.2, tmax=0.5, picks=picks,
reject=reject)

oneepoch=epochs['MI'].get_data()

#print(epochs)

#epochs['MI'].plot_psd(fmin=2, fmax=30)

#epochs.plot_psd_topomap(ch_type='eeg', normalize=True);

epochs['MI'].plot_psd_topomap(ch_type='eeg', normalize=True);

evoked = epochs['MI'].average()

#print(evoked)

#evoked.plot(time_unit='s')

#post MI data

info = mne.create_info(ch_names=ch_names, sfreq=sampling_freq, ch_types=ch_types);

#for trail in range(0,ntrail):

raw= mne.io.RawArray(eegdata[6,:,:], info) # reading neurofeedback 8 which correspon to
poststim nf 14 n14

raw.set_montage("standard_1020")

```

```

#raw.set_eeg_reference(['Cz'])
raw.filter(l_freq=8, h_freq=22)
#raw.plot_psd(fmax=50);
start, stop = raw.time_as_index([100, 115]) # 100 s to 115 s data segment
data, times = raw[:, start:stop]

plt.plot(raw.get_data()[0])

#data, times = raw[2:20:3, start:stop] # access underlying data
#plotraw.get_data()[0]
raw.plot(); # print raw data post
reject= dict( eeg=150e-6 ) # 150 µV

event_id = {'ready': 1, 'MI': 2, 'postrest': 3}

events = mne.find_events(raw, 'STI 001');
##fig = mne.viz.plot_events(events, event_id=event_id,
sfreq=raw.info['sfreq'],first_samp=raw.first_samp)

raw.info['bads'] += [ 'A1' ]
picks = mne.pick_types(raw.info, eeg=True, stim=False,exclude='bads')
epochs = mne.Epochs(raw, events, event_id, tmin=-0.2, tmax=0.5, picks=picks,
reject=reject)
epochs['MI'].plot_psd_topomap(ch_type='eeg', normalize=True);

```

```

twoepoch=epochs['MI'].get_data()

#TIME FREQUENCY ANALYSIS

# n_cycles = 2 # number of cycles in Morlet wavelet
# freqs = np.arange(7, 30, 3) # frequencies of interest
# from mne.time_frequency import tfr_morlet # noqa
# power, itc = tfr_morlet(epochs['MI'], freqs=freqs, n_cycles=n_cycles,return_itc=True,
decim=3, n_jobs=1)

# power.plot([power.ch_names.index('FC2')]); #TF graph for FC2

print(oneepoch.shape)

print(twoepoch.shape)

print(min([oneepoch.shape[0],twoepoch.shape[0]]))

minev=min([oneepoch.shape[0],twoepoch.shape[0]])

[ev,ch,sp]=oneepoch.shape;
preMI=np.zeros([ch,sp,ev])
#print(ev,ch,sp)
collectev=np.zeros([ev,sp])
for channel in range(0,ch):
    for evt in range(0,ev):
        collectev[evt,:]=oneepoch[evt,channel,:]
    preMI[channel,:,:]=collectev[:,:].T

#print(preMI.shape)

```

```

preMIre=preMI[:,:,0:mnev]
#print(preMIre.shape)

[ev,ch,sp]=twoepoch.shape;
postMI=np.zeros([ch,sp,ev])
#print(ev,ch,sp)
collectev=np.zeros([ev,sp])
for channel in range(0,ch):
    for evt in range(0,ev):
        collectev[evt,:]=twoepoch[evt,channel,:]
        postMI[channel,:,:]=collectev[:,:].T
postMIre=postMI[:,:,0:mnev]
print(postMIre.shape)
print(preMIre.shape)
def psd(trials):
    """
    Calculates for each trial the Power Spectral Density (PSD).

    Parameters
    -----
    trials : 3d-array (channels x samples x trials)

        The EEG signal

    Returns
    -----
    trial_PSD : 3d-array (channels x PSD x trials)

```

the PSD for each trial.

freqs : list of floats

The frequencies for which the PSD was computed (useful for plotting later)

'''

```
[nchannels,nsamples,ntrials] = trials.shape
```

```
trials_PSD = np.zeros((nchannels, 46, ntrials))
```

```
# Iterate over trials and channels
```

```
for trial in range(ntrials):
```

```
    for ch in range(nchannels):
```

```
        # Calculate the PSD
```

```
        (PSD, freqs) = mlab.psd(trials[ch,:,trial], NFFT=int(nsamples), Fs=Fs)
```

```
        trials_PSD[ch, :, trial] = PSD.ravel()
```

```
    return trials_PSD, freqs
```

```
cl_lab=['pre', 'post']
```

```
c11 = cl_lab[0]
```

```
c12 = cl_lab[1]
```

```
trials={ }
```

```
trials[c11]=preMIre
```

```
trials[c12]=postMIre
```

```

psd_r, freqs = psd(trials[c11])
psd_f, freqs = psd(trials[c12])
print(psd_f.shape)
print(freqs.shape)
trials_PSD = {c11: psd_f}
#print(trials_PSD.keys())
print(nchann)
plt.figure('pre')
for chan in range(0,nchann-2):
    plt.plot(np.mean(psd_r[chan,:,:] ,axis=1))
plt.figure('post')
for chan in range(0,nchann-2):
    plt.plot(np.mean(psd_f[chan,:,:] ,axis=1))
def plot_psd(trials_PSD, freqs, chan_ind, chan_lab=None, maxy=None):
    """
    Plots PSD data calculated with psd().

    Parameters
    -----
    trials : 3d-array
        The PSD data, as returned by psd()
    freqs : list of floats
        The frequencies for which the PSD is defined, as returned by psd()
    chan_ind : list of integers
        The indices of the channels to plot
    chan_lab : list of strings

```

```

    (optional) List of names for each channel
maxy : float
    (optional) Limit the y-axis to this value
'''
plt.figure(figsize=(12,5))

nchans = len(chan_ind)

# Maximum of 3 plots per row
nrows = np.ceil(nchans / 3)
ncols = min(3, nchans)

# Enumerate over the channels
for i,ch in enumerate(chan_ind):
    # Figure out which subplot to draw to
    plt.subplot(nrows,ncols,i+1)

    # Plot the PSD for each class
    for cl in trials.keys():
        plt.plot(freqs, np.mean(trials_PSD[cl][ch,:,:], axis=1), label=cl,linewidth=5)

# All plot decoration below...

plt.xlim(1,30)

if maxy != None:

```

```

plt.ylim(0,maxy)

plt.grid()

plt.xlabel('Frequency (Hz)')

if chan_lab == None:
    plt.title('Channel %d' % (ch+1))
else:
    plt.title(chan_lab[i])

plt.legend()

plt.tight_layout()

from numpy import linalg

def cov(trials):
    """ Calculate the covariance for each trial and return their average """
    [nchannels, nsamples, ntrials] = preMIre.shape
    covs = [ trials[:,:,i].dot(trials[:,:,i].T) / nsamples for i in range(ntrials) ]
    return np.mean(covs, axis=0)

def whitening(sigma):
    """ Calculate a whitening matrix for covariance matrix sigma. """
    U, l, _ = linalg.svd(sigma)

```

```
return U.dot( np.diag(1 ** -0.5) )
```

```
def csp(trials_r, trials_f):
```

```
    """
```

```
    Calculate the CSP transformation matrix W.
```

```
    arguments:
```

```
        trials_r - Array (channels x samples x trials) containing right hand movement trials
```

```
        trials_f - Array (channels x samples x trials) containing foot movement trials
```

```
    returns:
```

```
        Mixing matrix W
```

```
    """
```

```
    cov_r = cov(trials_r)
```

```
    cov_f = cov(trials_f)
```

```
    P = whitening(cov_r + cov_f)
```

```
    B, _, _ = linalg.svd( P.T.dot(cov_f).dot(P) )
```

```
    W = P.dot(B)
```

```
    return W
```

```
def apply_mix(W, trials):
```

```
    """ Apply a mixing matrix to each trial (basically multiply W with the EEG signal matrix)"""
```

```
    [nchannels, nsamples, ntrials] = preMIre.shape
```

```
    trials_csp = np.zeros((nchannels, nsamples, ntrials))
```

```
    for i in range(ntrials):
```

```
        trials_csp[:, :, i] = W.T.dot(trials[:, :, i])
```

```
    return trials_csp
```

```
W = csp(preMIre, postMI)
```

```
trials_csp = {c11: apply_mix(W, trials[c11]),
              c12: apply_mix(W, trials[c12])}
```

```
psd_r, freqs = psd(trials_csp[c11])
```

```
psd_f, freqs = psd(trials_csp[c12])
```

```
trials_PSD = {c11: psd_r, c12: psd_f}
```

```
plot_psd(trials_PSD, freqs, [0,11,-1], chan_lab=['CSP of MI', 'middle component', 'last component'], maxy=0.
```

## REFERENCES

- [1] Ahmed, Bilaal, et al. "Depression and anxiety: a snapshot of the situation in Pakistan." *International Journal of Neuroscience and Behavioral Science* 4.2 (2016): 32.
- [2] Mondini, Valeria, Anna Lisa Mangia, and Angelo Cappello. "Single-session tDCS over the dominant hemisphere affects contralateral spectral EEG power, but does not enhance neurofeedback-guided event-related desynchronization of the non-dominant hemisphere's sensorimotor rhythm." *PloS one* 13.3 (2018): e0193004.
- [3] J. Yang, H. Singh, E. L. Hines, et al., "Channel selection and classification of electroencephalogram signals: an artificial neural network and genetic algorithm-based approach," *Artificial Intelligence in Medicine*, vol. 55, no. 2, pp. 117–126, 2012.
- [4] F. S. Tyner, J. R. Knott. 1989. *Fundamentals of EEG technology, Volume 1: Basic concepts and methods*, Raven press, New York.
- [5] J. M. Aguilar, J. Castillo, and D. Elias, "EEG signals processing based on fractal dimension features and classified by neural network and support vector machine in motor imagery for a BCI," in *VI Latin American Congress on Biomedical Engineering CLAIB 2014*, Paraná, Argentina 29, 30 & 31 October 2014, A. Braidot and A. Hadad, Eds., vol. 49 of *IFMBE Proceedings*, pp. 615–618, Springer International Publishing, 2015.

- [6] M. Serdar Bascil, A. Y. Tesneli, and F. Temurtas, "Multichannel EEG signal feature extraction and pattern recognition on horizontal mental imagination task of 1-D cursor movement for the brain-computer interface," *Australasian Physical & Engineering Sciences in Medicine*, vol. 38, no. 2, pp. 229–239, 2015.
- [7] R. J. Williams and D. Zipser, "A learning algorithm for continually running fully recurrent neural networks," *Neural computation*, vol. 1, no. 2, pp. 270–280, 1989.
- [8] B. Blankertz et al., "The BCI competition III: Validating alternative approaches to actual BCI problems," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 14, no. 2, pp. 153–159, Jun. 2006.
- [9] [https://www.youtube.com/watch?v=XMizSSOejg0&ab\\_channel=JeremyMoeller](https://www.youtube.com/watch?v=XMizSSOejg0&ab_channel=JeremyMoeller)
- [10] Al-Kaysi, Alaa M., et al. "Predicting tDCS treatment outcomes of patients with major depressive disorder using automated EEG classification." *Journal of affective disorders* 208 (2017): 597-603
- [11] Nitsche, M.A. and W. Paulus, Excitability changes induced in the human motor cortex by weak transcranial direct current stimulation. *J Physiol*, 2000. 527: p. 633-9.
- [12] Turi, Z., W. Paulus, and A. Antal, Functional neuroimaging and transcranial electrical stimulation. *Clin EEG Neurosci*, 2012. 43(3): p. 200-8.
- [13] Nitsche, M.A., et al., Level of action of cathodal DC polarisation induced inhibition of the human motor cortex. *Clin Neurophysiol*, 2003. 114(4): p. 600-4.
- [14] Home-Based Transcranial Direct Current Stimulation Device Development: An Updated Protocol Used at Home in Healthy Subjects and Fibromyalgia Patients doi: 10.3791/57614 Published: July 14, 2018 Fabiana Carvalho<sup>1,2</sup>, Aline Patrícia Brietzke<sup>1,2</sup>, Assunta Gasparin<sup>1,2</sup>, Franciele Pereira dos Santos<sup>2,3</sup>, Rafael Vercelino<sup>4</sup>, Rafael Firmino Ballester<sup>2,3</sup>, Paulo Roberto Stefani Sanches<sup>5</sup>, Danton Pereira da Silva Jr<sup>5</sup>, Iraci L. S. Torres<sup>1,6</sup>, Felipe Fregni<sup>7,8</sup>, Wolnei Caumo<sup>1</sup>,
- [15] Fertonani, A. and C. Miniussi, Transcranial Electrical Stimulation: What We Know and Do Not Know About Mechanisms. *Neuroscientist*, 2016.
- [16] Boggio, P.S., et al., Effects of transcranial direct current stimulation on working memory in patients with Parkinson's disease. *J Neurol Sci*, 2006. 249(1): p. 31-8.

- [17] Liu, C.S., et al., Using transcranial direct current stimulation to treat symptoms in mild cognitive impairment and Alzheimer's disease. *Neurodegener Dis Manag*, 2017. 7(5): p. 17-329.
- [18] Woźniak-Kwaśniewska, Agata, et al. "Changes of oscillatory brain activity induced by repetitive transcranial magnetic stimulation of the left dorsolateral prefrontal cortex in healthy subjects." *Neuroimage* 88 (2014): 91-99.
- [19] "Predicting brain stimulation treatment outcomes of depressed patients through the classification of EEG oscillations." *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2016.
- [20] Ang, Kai Keng, et al. "Facilitating effects of transcranial direct current stimulation on motor imagery brain-computer interface with robotic feedback for stroke rehabilitation." *Archives of physical medicine and rehabilitation* 96.3 (2015): S79-