EEG-Based Detection of Directed Effective Connectivity and Microstate Analysis in Stress Disorder

Author

AKASHA AMAN

00000328106

Supervisor

DR. AHMAD RAUF SUBHANI

DEPARTMENT OF ELECTRICAL ENGINEERING

COLLEGE OF ELECTRICAL & MECHANICAL ENGINEERING

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY

ISLAMABAD

October 2024

THESIS ACCEPTANCE CERTIFICATE

It is certified that final copy of MS/MPhil thesis written by Ms. Akasha Aman (Registration No. 00000328106) Entry-2020, of (College of E&ME) has been vetted by the undersigned, found complete in all respects as per NUST Statutes/Regulations, is free of plagiarism, errors, and mistake and is accepted as partial fulfillment for award of MS degree. It is further certified that necessary amendments as pointed out by GEC member of the scholar have also been incorporated in the said thesis.

Signature: ____________

Name of Supervisor Dr. Ahmad Rauf Subhani

Date: 22-oct-2024

Signature (HoD): ____________

Dr. Qasim Umar Khan

Date: 22-oct-2024

Signature (Dean) ____________

Brig Dr. Nasir Rashid

Date: 22 OCT 2024

Dedication

I dedicate this research work to my esteemed Parents, my respected supervisor and my honorable teachers who teaches, supports and motivate in each and every step of life and who have inspired to do great things and take risks which I have never tried.

Acknowledgement

I have great privilege and reasons to express my deepest bones of gratitude and here felt respect to all persons who help us. First and foremost, I would like to express my deepest gratitude and humblest thanks to ALLAH ALMIGHTY, who is the most Merciful and Beneficial of all, who presented us, the capacity to finish this, anticipate. I express my sincere gratitude and indebtedness to Dr. Ahmad Rauf Subhani who provide opportunity to work under their supervision. He supported throughout with their patience and great knowledge. I would like to extend my heartfelt gratitude to my parents. Their sincere prayers for my success have always been highly appreciated and have been the source of motivation, courage and strength. I also place on record, our sense of appreciation to every last one, who straightforwardly or by implication, has lent a helping hand. Finally, I wish to acknowledge that in our daily work I have been blessed to meet with a friendly and cheerful group of fellow students.

Abstract

Stress is a common phenomenon that affects individuals from all walks of life and it is associated with physical and physiological illness. Brain plays a key role in determining stress. This study represents the comprehensive framework for analyzing EEG data to investigate the connectivity and network dynamics under stress and normal. Our primary interest is how classifying stress in network so this is assessed through effective connectivity in which we use partial directed coherence to measure how one region influences other. Connectivity networks are computed using adjacency matrix over time and subjected to graph theoretical analysis to focus on degree and betweenness Degree highlighted highly connected region and betweenness revealed brain central region that facilitate global communication across different brain region. Further analysis in microstate that demonstrate the distinct patterns in brain, from this analysis microstate 4 show high duration and transition in normal conditions, while microstate 5 show high duration and transition in stress condition. Microstate features are used for classification into stress and normal by using three machine learning classifiers: random forest (RF), K-nearest neighbors (KNN) and support vector machine (SVM). In the classification analysis, Random Forest achieved 94% accuracy followed by KNN and SVM. Furthermost, this approach provides valuable insights into brain connectivity and demonstrate the utility of microstate features in stress classification.

Keywords: Stress, Independent Component Analysis, Electroencephalogram (EEG), Effective Connectivity, Partial Directed Coherence(PDC),Graph Theoretical Analysis, Microstate Analysis.

Table of Contents

List of figures

List of Table

List of Abbreviations

EEG Electroencephalography

ML Machine Learning

FC Functional Connectivity

MSC Mean Squared Coherence

FIR Finite Impulse Response

ASR Artefact Subspace Reconstruction

LR Learning Rate

SVM Support Vector Machine

KNN K-Nearest Neighbors

PET Positron Emission Tomography

fMRI Functional Magnetic Resonance Imaging

ICA Independent Component Analysis

TI Time Invariant

ST Short Time

IIR Infinite Impulse Response

Chapter 1: Introduction

1.1 Stress

Stress gained widespread recognition in its biological context through Hans Selye, who described it as "the non-specific response of the body to any demand for change" [1]. Stress is a physiological and psychological response to pressure or demand from the environment. Stress is a process that strains an individual capacity, impacting both physical and mental abilities. It is a risk factor for numerous health issues, including heart attack, stroke, hypertension, and even sudden death [2]. Additionally, stress is associated with decision-making, performance, and learning. It can stem from both positive and negative emotions [3]. However, when stress is prolonged and associated with negative emotions, it can lead to more serious emotional states such as depression, sadness and anger.

According to World health organization, stress impacts both the body and the mind. Excessive stress can lead to physical and mental health issues, as 3.6% of the global population experienced stress [4].

Stress is further classified into three types; acute stress, chronic stress and episodic stress [2]. Acute Stress is a natural and adaptive response to immediate challenges, typically involving short-term stress and this often occurs in response to specific events or situations. Next is Chronic Stress a long-term form of stress that continues over an extended period. lastly, Episodic Stress arises when stressors occur more frequently for a limited duration [2].

1.2 Importance of Stress Detection

Stress trigger behavioral and mental changes due to lack of relief from various challenges. Stress can arise from various sources. Here are some common causes of stress and these include: work pressure, financial concerns, health concerns, daily hassles and traumatic events.

Stress can affect the body in variety of ways. These include the nervous system, endocrine system, musculoskeletal system, digestive system and mental health further stress symptoms are indicated by emotional and physical symptoms. Emotional symptoms include anxiety, depression, anger, moodiness, loneliness and forgetfulness whereas physical symptoms include headache, high blood pressure, muscle aches, sweating, cramps, fainting and sleeping difficulties.

These symptoms also alter the behavior of individuals experiencing stress, such as sudden energy outbursts, food craving and frequent crying are important symptoms of stress [4]. Prolonged stress can lead to hypertension and contribute to conditions such as bowel disease and cardiovascular disorders [5]. Therefore, researchers focus on detecting stress at its early stages to prevent the onset of severe health complications.

Stress detection is important for individuals to recognize the impact of excessive stress on their well-being. The research community is concentrating on identifying the most effective methods for early detection of stress.

1.3 Methods For Stress Detection

Stress is now recognized as a societal issue. Stress assessment is typically conducted using two primary measures i.e. Subjective and objective. In subjective measure self-report scales, interviews and journals are used for stress detection but these are not reliable methods to detect stress. On the other hand, objective measure is further classified into physiological, neurological and behavioral. Physiological and neurological techniques are also mentioned in table 1 and 2. Heart rate is one of the measures for stress using ECG. As we know stress is linked with brain so this can be measure by hormones. The human cerebrum is regarded as the main source of stress. Cortisol is the key hormone for body response to measure stress. Stress is also detected on the basis of sleep pattern and social behavior. When someone is disturbed due to their sleep cycle and the underlying cause can be stress. This is not the correct way to detect stress because there is multiple reason behind the disturbed sleep and social behavior. So, the method which is discuss in this study to measure stress is Electroencephalography (EEG). Two types of responses i.e. physical and physiological measures are outlined in tables I and II below [6].

Table 1 Physical Methods for Stress Measurement

S. No	Technology Name	Physical Techniques
1.	Infrared Eye Tracking	Eye Activity
2.	Automated Gesture Analysis	Body Gesture

Table 2 Physiological Methods for Stress Measurement

S. No	Technology Name	Symbol	Physiological Techniques
1.	Electrocardiography	ECG	Electrical activity of the heart

2.	Electroencephalography	EEG	Electrical activity of the brain
3.	Electrodermal Activity	EDA	Skin Reaction
4.	Galvanic Skin Response	GSR	

Both physical and physiological methods form the basis of traditional stress detection systems. The major drawback of physical technologies for measuring stress is affected by external factors, including room temperature, anxiety and sweating. The brain waves that humans produce is a reflection of these elements.

According to research [6], stress can be identified using physiological traits of humans derived from physiological signals. The features that are physiological in nature, manifesting as electrical activity in the brain, serve as the basis for stress detection. In the next, we discuss about electroencephalography, how we used this for stress detection.

1.4 Background

In the last twenty years, multiple studies have been conducted to detect stress in humans using physiological measure. Some of the studies discuss the techniques which are used for stress detection. These techniques are further classified into various methods for detecting stress, which can be quantified through surveys, questionnaire or the monitoring of individuals to assess changes in physiological signals. In real-time applications, these signals can be measured and analyzed to evaluate stress levels [7], which are categories into two types:

I. Invasive methods
II. Non- invasive methods

These methods are further divide into two types:

I. Electrocorticography (ECoG)
II. Local Field Potential (LFP)

Both methods are recognized for delivering high precision and resolution along both spatial and time axes. Research indicates that electrocorticography (ECoG) involves placing electrode pads directly on the brain's surface to record brain waves. This technique achieves high resolution and wide bandwidth; however, it is limited by the need for surgical intervention to position the electrodes [7].

Some techniques have drawback of involving invasive procedures like hormone analysis this highlights the necessity for non-invasive, effective, reliable and accurate methods [8]. Stress can be measure from physiological signals which used different methods from non-invasive type [6],[7],[8].

Table 3 Comparison of different non-invasive measures

S. No	Non-Invasive Measures	Advantage	Limitations
1.	Functional Magnetic Resonance Imaging (fMRI)	Moderate to high spatial resolution	The temporal response of blood flow is gradual.
2.	Magnetic Encephalography (MEG)	High resolution	Very expensive
3.	Positron Emission Tomography (PET)	High resolution	Limited temporal resolution and considerable safety restrictions due to radiotracers
4.	Blood Pressure (BP)	Ease of use, measurement sensitivity, and availability of normotensive data.	The absence of prospective mortality data makes it unsuitable for determining whether treatment is required.
5.	Blood Volume Pulses (BVP)	Non-intrusiveness and affordability	The requirement for individual calibration and its tendency to drift over short periods.
6.	Galvanic Skin Resistance (GSR)	Affordability and immediate accessibility	Diurnal fluctuations affect the timing of assessments, influencing the results.
7.	Electromyography (EMG)	The most accurate, suitable, and dependable instrument for calculating muscle activity	Decreased clinical yield in certain instances, with technical challenges from obesity, advanced age, and signal interpretation complexity.
8.	Electrocardiography (ECG)	Equipment is widely available, and its accuracy	Lower sensitivity compared to alternative stress imaging

		has been confirmed across various populations	techniques and low specificity.
9.	Electroencephalography (EEG)	Exceptional temporal resolution, with no significant safety concerns	Low spatial resolution

EEG is the preferred option for surpassing ECG, EMG, PET and fMRI in providing accessible and reasonably priced insights into brain activities with high temporal resolution. It can be concluded that EEG is an excellent instrument, as it utilizes a non-invasive technique to receive feedback from stress hormones, making it both accurate and reliable for stress detection.

1.5 Electroencephalogram (EEG)

Brain is the primary organ that reacts to various feelings and emotions [9]. The brain controls both direct and indirect electrical signals that flow through the human body, acting as the central nervous system of living things. An electroencephalogram (EEG), a medical procedure, is used to record this electrical activity that neurons produce [10]. Electric currents between the brain cells in the cerebral cortex generate an EEG signal. Small metal disc connected to thin wires, known as electrodes, are placed on the scalp to facilitate the transmission of signals to a device for data recording. These electrical changes achieve high temporal resolution on the order of milliseconds or even microseconds. [10], [11]. EEG recordings are represented as graphs that illustrate the electrical activity produced by the brain over time.

By measuring the potential difference between two electrodes spaced apart and recording the total potential of the neurons, an EEG channel is formed. The 10-20 Standard electrode setup [10] is used to measure EEG, as shown in Fig. 1.

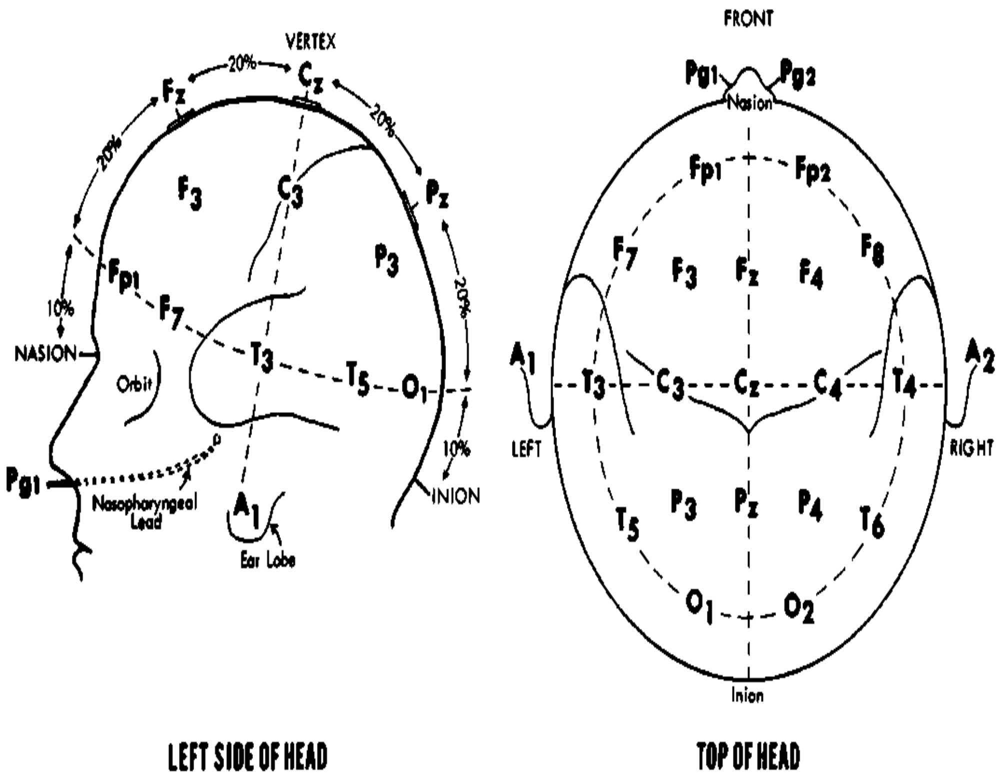

Figure 1. 10-20 Standard electrode system

EEG is a useful electrophysiological tool for early detection, assessment, and treatment because it is a potent method for tracking the nonlinear electrical activity of the brain's nerve cells.

1.6 Research Gap

To date, the exploration of EEG through effective connectivity for the classification of stress remains largely unexplored. Additionally, the analysis and application of microstate analysis in EEG data, specifically for stress classification using connectivity measures, represents an uncharted area in the field. These gaps highlight the need for further research, as combining effective connectivity with microstate analysis could provide deeper insights into brain dynamics and help improve the accuracy of stress detection using EEG.

1.7 Objectives

- Detection of stress using EEG data identify and investigate different patterns aiming to understand the neural connections associated with stress in terms of both dynamic connectivity and microstate analysis

- Assess the sensitivity and specificity of PDC in capturing stress-related alterations in brain connectivity. Examine changes in the occurrence, duration, and topography of microstates during stress-inducing tasks or scenarios

1.8 Structure

Chapter 1: This section is comprising of the Introduction of Stress and EEG and objectives of the proposed system.

Chapter 2: This section is about literature review where we further discuss the methods by which we detect stress

Chapter 3: This section comprises the major tasks i.e. the methodology of research

Chapter 4: This segment shows the results of the implemented work.

Chapter 5: This section contains conclusion.

References: This section contains the references of the articles and the books we have cited.

Chapter 2: Literature Review

The thorough summary highlights the international efforts undertaken so far in the field of stress reduction and deduction methods. It looks at the gathering and evaluation of data that is mostly physiological in nature and recognizes the connection between different physiological characteristics and felt stress, which was the main emphasis of earlier studies on stress assessment.

Recent advancements in neuroimaging technology have facilitated the use of electroencephalography (EEG) for measuring and recording the electrical activity of the human brain. This technique represents a significant advancement in the field of neuroscience, aiding neurologist [12]. Variations in brain wave patterns are crucial for diagnosing mental disorders, including stress.

According to Saidatul, A et.al. [13] The two main parts of the EEG stress assessment approach are feature extraction and stress categorization. EEG features are categories into three types: synchronicity-domain, time-domain, and frequency-domain features. By analyzing amplitude in connection to energy, variability, coefficient of variation, and features associated with higher-order crossings, time-domain features concentrate on temporal information. Conversely, the most commonly utilized frequency-domain features derived from EEG signals include delta (0.5-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), beta (14-30 Hz) and gamma (30-50 Hz) bands. These frequency bands encompass information pertinent to stress and psychological disorders.

The authors of [14] discuss in papers that EEG offers high temporal resolution, making it a valuable tool for examining fluctuations in mental states. It is used to capture brainwave patterns that correspond to stress responses, linking specific EEG bands to different stress levels. Due to their cost-effectiveness and non-invasiveness, they make an EEG a valuable tool for stress detection.

EEG signals encompass information generated by a intricate and densely connected network of neurons. Brain connectivity provide us how brain areas interconnect with each other. Brain connectivity can be categorized into neuroanatomical (or structural), functional, and effective connectivity, as noted by Horwitz B. et.al. [15]. EEG does not provide direct insight into structural connections; instead, it is used to estimate functional and effective connectivity. The relationships between various brain regions are referred to as functional connectivity, and they are shown by the temporal coherence within networks. The methods employed to determine functional connectivity can yield varying conclusions, influenced by elements like the kind of stressor, the number and arrangements of electrodes, and the degree of connection between brain units. This variability can occur even within data from the similar modality or when using the same task [16].

The authors of [17] mentioned that effective connectivity, refers to the causal influence that one neural region exerts over others. It is determined by integrating imaging techniques such as EEG and magnetoencephalography (MEG) with mathematical models of interconnected brain regions. Effective connectivity represents the most basic circuit that outlines the connection between two neurons. It illustrates how neural networks influence one another [18]. Effective connectivity, unlike the non-directional and correlational nature of functional connectivity, evaluates the directed influences between different brain regions. Connectivity metrics that are both efficient and functional are crucial for understanding brain behavior in both stressed and non-stressed conditions. Various features are available to detect these connectivity measures [19].

Xi et al. [20] this study investigated Coherence analysis, which seeks to determine the functional relationships among various brain areas. This analysis is conducted by examining the amplitude and phase of signals recorded from EEG electrodes. The results indicated a significant increase in coherence across all frequency bands during multilevel stress assessments.

Al-Shargie et al. [21] examined that Magnitude Square Coherence (MSC) functions as a supplementary indicator of functional connectivity associated with stress. A significant reduction in functional connectivity is observed when comparing control conditions to stress situations within the intra-hemispheric prefrontal cortex (PFC). The EEG connectivity maps showed that when sleep deprivation was utilized as a stressor, there was a rise in beta coherence throughout the entire scalp and a reduction in MSC for the alpha band across the anterior scalp region.

Zanetti M. et al. [22] Mutual information (MI) is utilized to assess the resemblance of the jointly distributed probability between two EEG signals. During stress, MI is reflected in maps of EEG connectivity. According to the study in [23], MI did not exhibit a notable increase in the EEG map during Stroop task. However, under the physical strain of lack of sleep, the linear area saw a significant declines, whereas the nonlinear area in the head's anterior, central, and temporoparietal regions showed a considerable rise.

Phase lag is utilized to determine the time difference between two EEG readings coming from distinct brain areas. Al-Shargie et al. [24] demonstrated that the phase lag technique plays a significant role in distinguishing between the conditions of control and stress across various levels. Its primary drawback, though, is that it is impacted by the volume conduction problem and fails to reflect the directionality of connectivity.

The Phase-slope index (PSI) is a metric for assessing phase synchronization that is unaffected by volume conduction or shared reference influences. Research in [25] shows that chronic stress can alter various connectivity patterns among brain regions as one is observing outside stimuli. Darzi et al. [26] demonstrated high performance using PSI features. In contrast, Khosrow Abadi et al. [27] reported lower accuracy for the Phase-Slope Index (PSI) compared to the Partial Directed Coherence (PDC) features. Additionally, PSI may not accurately represent the directionality of EEG signals.

Partial directed coherence (PDC) is employed to determine the weight and direction of the information flow between multivariate frequency domain datasets. By utilizing multivariate analysis, PDC captures the phenomena of stress while preserving the complete information from data involving multiple variables. Specifically, PDC can be used to anticipate two distinct coherences from the classical coherence function. Granger causality and Akaike information criterion are two of the determined parameters that affect the direction of flow within a specific frequency range across two channels. According to Al-Shargie et al. [28] the functional coupling diminishes as the amount of tiredness rises under stress. To regulate the EEG multichannel analysis's negative causality, Generalized Partial Directed Coherence (GPDC), is a crucial PDC version to provide functional connectivity assessment. GPDC traits were employed by Khosrowabadi et al. [29] to identify stress and non-stress instances.

Al Ezzi A et al. [30] shown that using PDC in conjunction with graph theoretical measures facilitates comparison with healthy controls (HCs) and aids in estimating the severity of SAD. This combination enhances the ability to detect key features of effective brain networks. A graph theoretical measures discuss according to functional connectivity. Xefteris VR et al. [31] this paper suggests an approach for combining peripheral physiological information with EEG analysis. The novelty of this research stems from the application of graph theory quantifies EEG data for classification of arousal and valence. The primary objectives of the analysis are to compute EEG functional connectivity networks and extract features based on graph theory from them. Accuracy is the main concern in this study and the proposed methodology is not reliable.

Kim, K et al. [32] shown a demonstration on the usefulness of EEG microstate features in evaluating cognitive performance but the limitation here is the limitation of dataset according to their proposed methodology. Al-Ezzi A et al. [30] found that the impact of cognitive reevaluation in conjunction with upbeat music for controlling emotions was examined by examining the functional connectivity of microstates based on EEG. This analysis is on electrophysiological method and these are not generalizable.

Bhatnagar, S et al. [33] in this study EEG microstates identify as biomarkers across multiple mental health disorders and treatment interventions. And the limitation is interpretability of features.

Christoph M. Michel et al. [34] The use of EEG microstate methods in clinical and experimental research is growing, with well-established analysis procedures and objective quantifiers. However, some key challenges remain unresolved in this study, such as determining the optimal number of microstate and the methods for defining them.

From above studies they conclude that they explored EEG connectivity to understand stress and brain dynamics, employing methods like coherence analysis, mutual information, phase lag, and partial directed coherence (PDC). While coherence and mutual information show significant changes under stress, phase lag and phase-slope index (PSI) have limitations in capturing directionality. PDC, combined with graph theoretical measures, provides valuable insights into stress-related connectivity, though some methodologies still require refinement. EEG microstate analysis shows potential in identifying cognitive and mental health markers, but challenges remain in interpretability and dataset limitations. Overall, further research is needed to combine microstate features with connectivity measures like effective connectivity to improve accuracy and generalizability.

Chapter 3: Methodology

This chapter outlines the methods employed to identify pertinent literature for the study. It consists of several sections, as depicted in the figure below and described here:

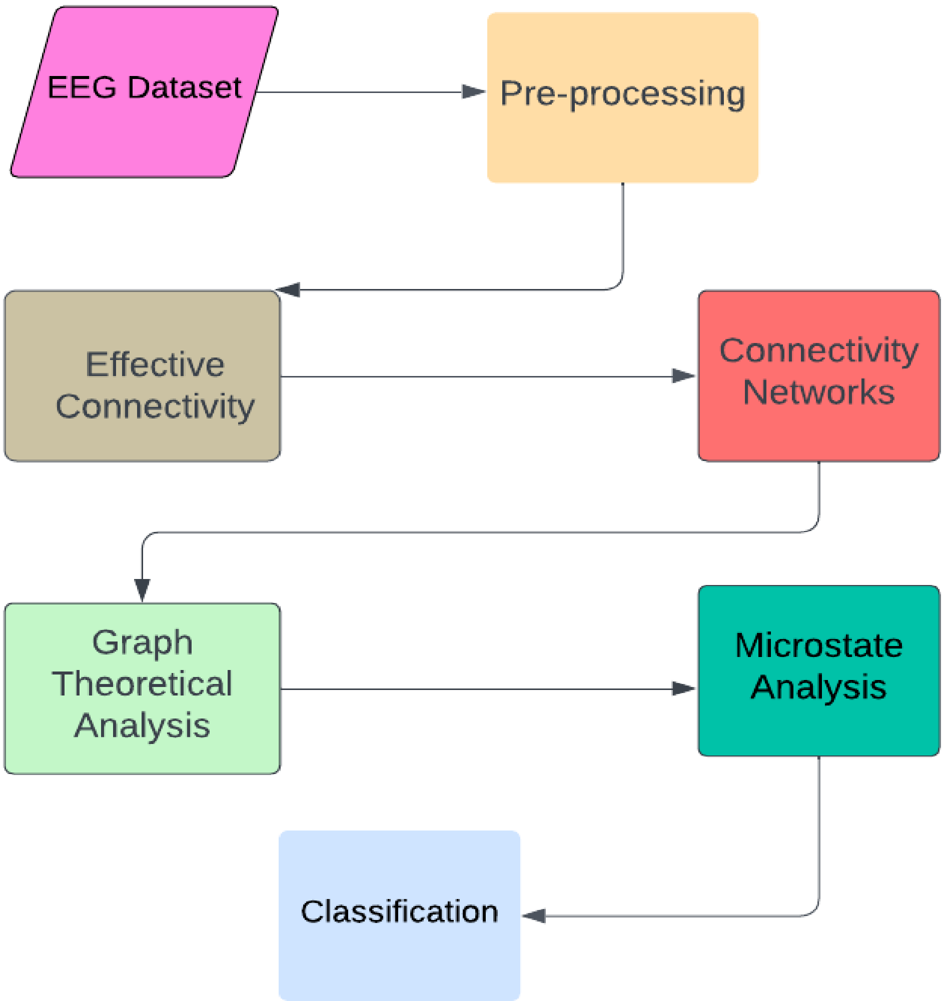

Figure 2 A schematic representation of the process for constructing an effective connectivity network from EEG data utilizing graph theory

3.1 EEG Data Acquisition

EEG information was obtained using the eegosports amplifier (ANT Neuro, Enschede, Netherlands) with 32 gel-based Ag/AgCl arranged on an EEG cap following the Extended International 10-20 system. The electrodes grounded at AFz and referenced to CPz, as per manufacturer's guidelines (eemagine Medical Imaging Solutions GmbH, Berlin, Germany), eliminating the need for re-referencing the EEG data. All electrode reading were kept below 10kΩ during the recording sessions.

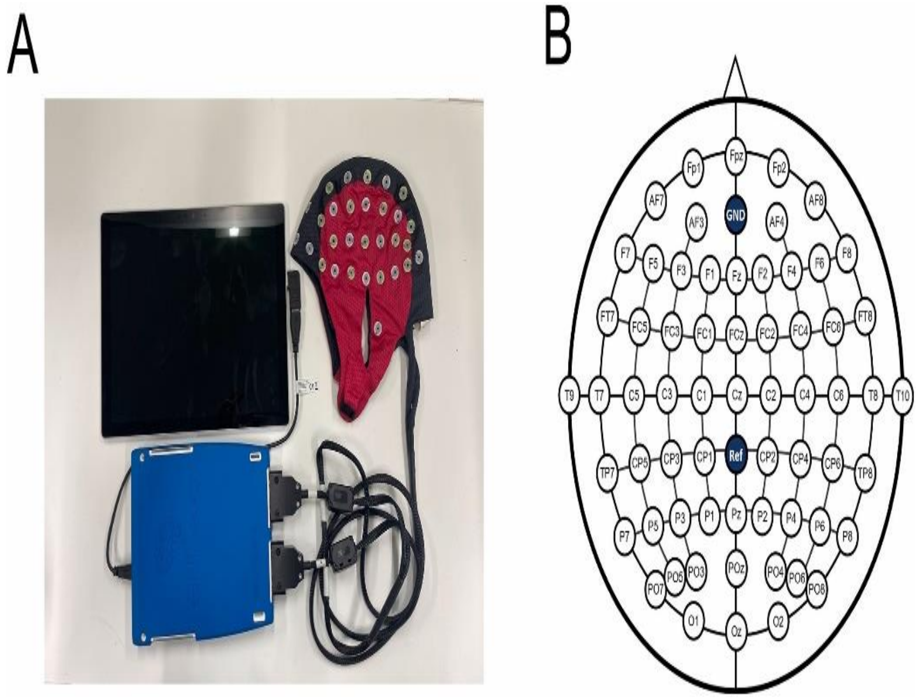

Figure 3 Image of EEG device and electrode position schema.

A) EEG device (eego™sports; ANT Neuro, Enschede, Netherlands) used in this study and (B) the electrode position schema following the 10-20 system.

In above figure, EEG equipment and electrode placement scheme is shown. So here is the representation of how these electrodes are shown with according to the brain lobe. Lobes of the brain are frontal, parietal, temporal and occipital and in 10-20 standard placement of channel location is shown in figure 3.3

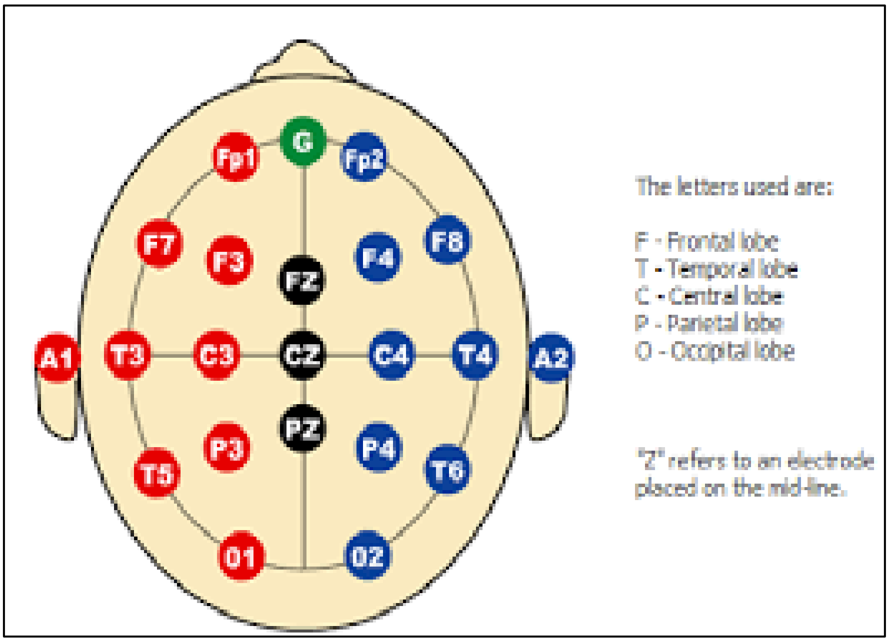

Figure 4 Channel Location according to 10-20 standard

3.2 Data Preprocessing

EEG preprocessing is performed by following procedure which is shown in above Fig 3.1. Raw EEG data files stored in .cnt (Continuous) format, imported using EEGLAB toolbox in MATLAB. These files are further categories according to gender Male and Female, and after this divide stress into two levels, stress and normal. In this study, dataset contains electroencephalography (EEG) recordings for 36 subjects, 18 females and 18 males.

The dataset contains both raw data and cleaned data filtered. To preprocess the raw EEG data, independent component analysis (ICA) and bandpass filtering were implemented. The data originally contained 64 channels, which are reduced to 35 relevant channels to focus on brain-activity and exclude non-essential data. A causal band pass filter is applied using EEGLAB legacy finite impulse response (FIR)filter with 1Hz to 50Hz bandwidth to retain the EEG frequencies of interest while removing unwanted low-frequency drift and high frequency noise. The causal filter ensured that phase delays are not compensated for preserving the causal relationships in the data which is crucial for effective connectivity and causal analysis. Fast-ICA is employed to remove noise cause by eye blinking artifacts, muscle artifacts and heart rate. Subsequent analyses were conducted using custom MATLAB scripts and open-source toolboxes, including the (1) Brain Connectivity Toolbox for graph theoretical analysis and (2) EEGLAB for generating topographical maps [35]. Time domain analysis is implicitly performed as a part of pre-processing step, where the EEG data is segmented and cleaned for further analysis. The frequency domain analysis is not performed so short-time Fourier transform is applied to find out which frequency band is affected by stress.

3.3 Frequency Domain Analysis

When analyzing in the frequency domain, EEG data are converted using techniques such as Fourier Transform (FFT) to extract power across different frequency bands. The five frequency bands that make up an EEG signal are Delta (1-4Hz), Theta (4-7Hz), Alpha(8-15Hz), Beta(16-31Hz) and Gamma (>32 Hz). In our study, we apply this to find out which frequency band is affected by stress.

3.4 Fast Fourier Transform

A mathematical technique called the Fast Fourier transform is employed to analyze and represent data in terms of time, which is then changing within the frequency domain. In the context of EEG signals, FFT is utilized to convert signals from the frequency domain to the time domain. This transformation helps in identifying and analyzing the different frequency components that associated with different mental conditions [36].

The Discrete Fourier Transform (DFT)can be computed efficiently using the FFT technique. As a result, O (N^2) computational complexity is reduced to *O (N log N)*. The FFT algorithm pertains to the same formula as DFT but optimizes the computation the breaking down the DFT.

3.5 Power Spectral Density

After applying FFT, we compute power spectral density to analyze the power of different power frequency bands. The PSD *P(f)* is computed as:

$$P(f) = \frac{1}{N} |X(f)|^2 \tag{1}$$

Where:

P(f) represents the power of signal strength at frequency f.

X(f) denotes the signal's Fourier transform at frequency f.

N indicates the signal's sample count.

In this study, the application of FFT helps in study of the EEG signals' frequency components, which is helpful for identifying that brain activity changes in this case i.e. stress. It allows to break down complex time-domain signals into understandable frequency components to analyze brain rhythms [35]. In fig it shows how a signal is decomposed into FFT.

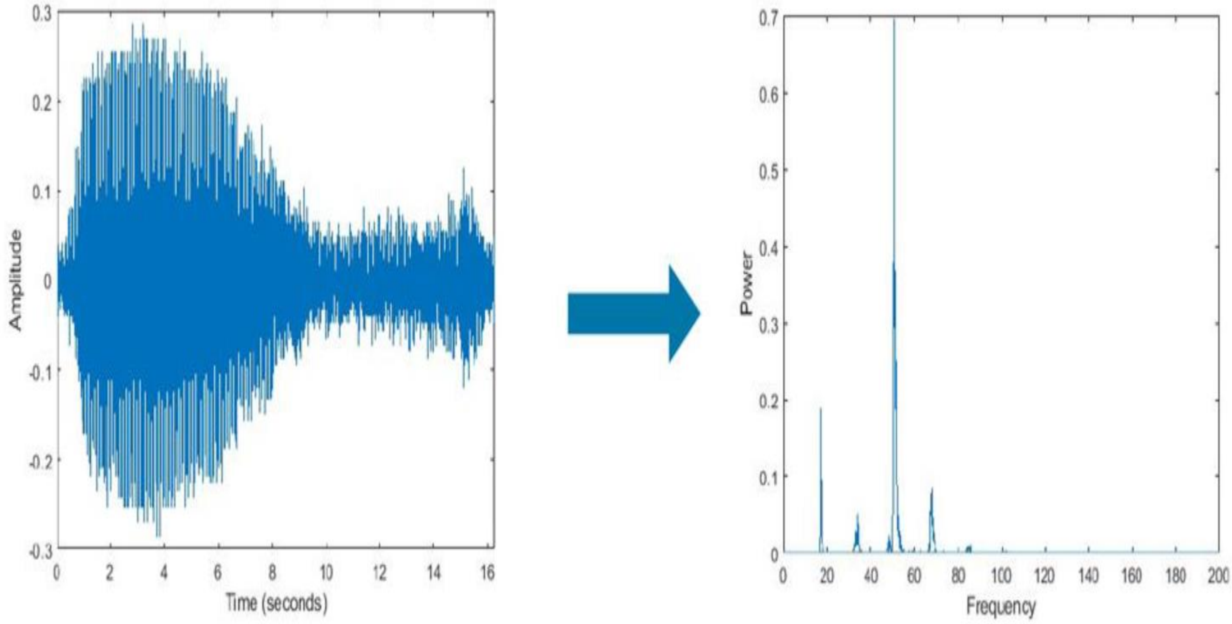

Figure 5 DFT decompose into FFT.

3.6 Effective Connectivity

Effective Connectivity analysis focuses on understanding the directional impact one neural region has on another, providing insights into the causal relationships among various brain regions [36]. This is particularly useful in EEG studies related to cognitive processes like stress detection, where understanding the flow of information across brain regions. Effective connectivity reveals how one region influences activity on another, rather than simply identifying correlations or statistical dependencies like in functional connectivity [37]. Effective connectivity can help the neural mechanism behind stress responses by showing how brain regions interact dynamically over time. So, this information is found out by brain lobes on the basis of effective connectivity. From the four lobes of brain, we can get information that which lobe of the brain causes stress. These are frontal, parietal, temporal and occipital. In Fig. 4.2 Lobes of the brain are shown.

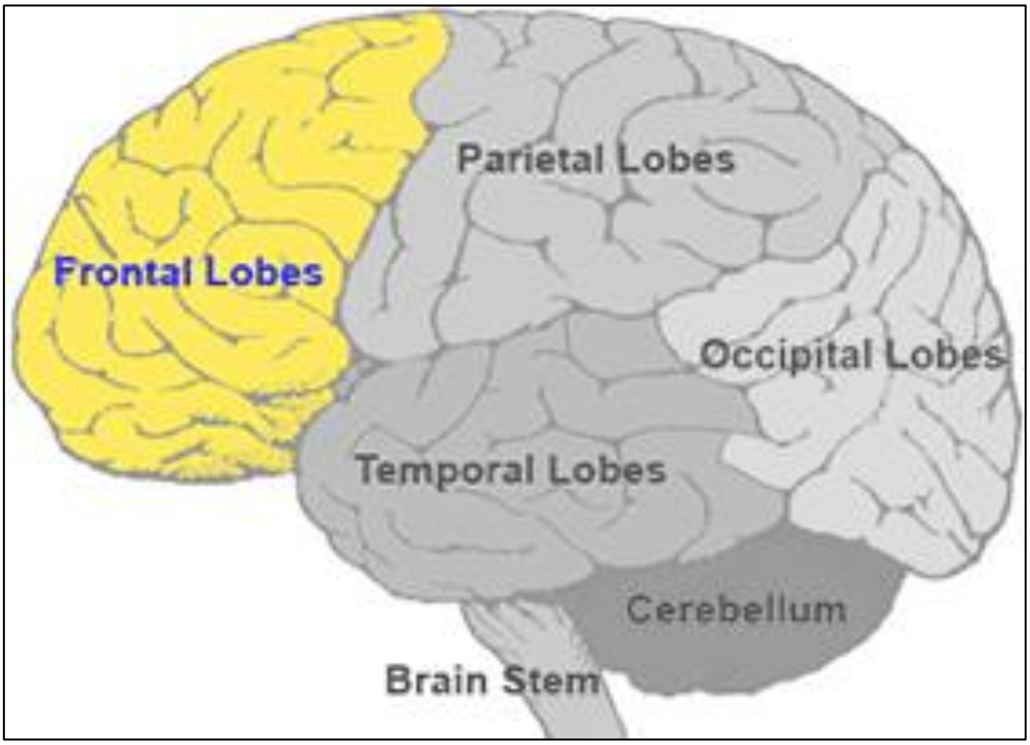

Figure 6 Lobes of the brain

3.7 Partial Directed Coherence (PDC)

Partial Directed Coherence (PDC) is a frequency-domain measure used to assess effective connectivity and is widely used in EEG [36]. PDC provides a way to measure the directed influence between brain signals, indicating whether the activity of one part of the brain influences that of another, which is a core aspect of effective connectivity. PDC works in the frequency domain to analyze how brain regions interact over different frequency brands (e.g. alpha, beta, gamma). PDC is applied to multivariate EEG signals and is well-suited to handle complex brain networks with multiple interacting regions. Compared to other methods, PDC tends to reduce false interactions or cross-talk between signals, making the analysis of brain connectivity more reliable. PDC is robust to noise and outliers, which makes it suitable for analyzing real EEG data. Using the approach described by Baccala and Sameshima [38] as a spectral measure with several variables for identifying the causal relationships that are directed within a multivariate data collection between a certain pair of time series signals, the PDC data were utilized in order to evaluate the connectivity.

PDC is a technique that measures the relationship between two signals out of multiple n signals, accounting for the influence of other signals and mitigating the effects of volume conduction. It enhances the concept of partial coherence by measuring directional influence and is calculating using a Multivariate Autoregressive (MVAR) model.

X(t) represents a set of estimated signals from *m* channels, and the order *p* model is represented by:

$$X(t) = \sum_{r=1}^{p} A(r)X(t-r) + E(t) \tag{2}$$

Where *m* denotes the number of nodes and $X(t)$ is the signal vector with *m* channels at time *t*. *E(t)* embodies the estimated error and *A(r)* is the autoregressive coefficient's coefficient matrix at lag *r*. The model order *p,* which was calculated using AsympPDC [39] package to reduce the Akaike's information criterion (AIC) value, determines the fitting outcomes of the MVAR model.

The frequency domain representation of *A(f)* can be defined by after obtaining the coefficient matrix of the MVAR model containing the values of *A(r)*.

$$A(f) = \sum_{r=1}^{p} A(r)\, e^{-i2\pi fr} \tag{3}$$

Next, the discrete Fourier transform is used to the autoregressive coefficient matrix *A(r)*:

$$\overline{A}(f) = I - A(f) = [\overline{a}_1(f)\overline{a}_2(f)\cdots\overline{a}_m(f)] \tag{4}$$

where, $\overline{a}_k(f)$ is the *k*-th column of the matrix $\overline{A}(f)$,and the elements of $\overline{A}(f)$ are defined as:

$$\overline{A}_{lk}(f) = \{ \begin{matrix} 1 - \sum_{r=1}^{p} a_{lk}(r)\, e^{-i2\pi fr}, & \text{if} l = k \\ - \sum_{r=1}^{p} a_{lk}(r)\, e^{-i2\pi fr}, & \text{otherwise} \end{matrix} \tag{5}$$

Where, the node index is indicated by *l* and *k*.

PDC value, which can be described as follows, indicates the direction and strength of the flow of information at frequency between nodes *k* and *l* at frequency *f,* which is defined as:

$$PDC_{lk}(f) = \frac{\overline{A}_{lk}(f)}{\sqrt{\overline{a}_k^H(f)\overline{a}_k(f)}}, l \neq k \tag{6}$$

Where, H represents the transposition of a matrix.

The adjacency matrix was generated by figuring out a pair of nodes' PDC values, followed by threshold processing. In this step, if the PDC value exceed a threshold T, it indicates a connection amongst the nodes; conversely, when the PDC value falls below T, no connectivity is inferred. This thresholding technique is commonly employed to eliminate false relationships as well as and create the brain network, a sparse connectivity matrix [40]. Variations in the threshold T of the adjacency matrix can significantly influence the numerical properties and the causal brain network's topological configuration.

3.8 Adjacency Matrices

With the values of PDC between each paired association for each EEG sub-band (delta, theta, alpha and beta) for stress, the connectivity matrices are constructed. Transforming the connection matrices into adjacency matrices involved applying a threshold.

In this study, selecting an appropriate threshold is critical for creating a matrix of adjacency based on the association matrix. A significance level-based approach is applied to establish the threshold T. Subsequently, the effective network is transformed into an unweighted graph with direction, wherein the connection matrix, comprised of PDC values for every pair of directed nodes is transformed into an adjacency matrix A using the threshold T. In this network, the strongest interactions are indicated by the threshold T, which can be understood as the ratio of the number of real effective connections to the total number of connections available. To investigate the characteristics of the efficient connectivity networks at varying connection strengths, the threshold T is set within the range of 0.1 to 0.9 with increments of 0.05.

For each threshold value, the adjacency matrix derived from PDC are filtered to retain only connections stronger than the threshold. So, the resulting connectivity matrices are then used to examine network metrics, such as local efficiency. After computing matrices, the topographical properties of the adjacency matrix are quantified using graph theoretical analysis.

3.9 Graph Theoretical Analysis

Graph theory is a crucial tool for analyzing electrical circuits and chemical structures. Its contemporary evolution started with the development of the scale-free network model in the late 1990s [41], which facilitated the understanding of brain connectivity patterns. During the past two decades, the graph theory has been applied to quantifies neurophysiological data gained significant interests in the field of biology and neuroscience to diagnose brain disorder like stress. One of the objectives of the current study is to bring attention to cognitive neuroscience.

Topographical connections patterns are quantified by graph theory, which is then applied to brain connectivity networks. The data are shown as a Graph (G), a simple topographical representation composed of a set of V vertices (nodes) joined by edges (E) that are linkages, where G = (V, E) as shown in Fig.3.3, to help with better understanding of connectivity patterns. The human brain network is organized on a large scale is represented by nodes, which correspond to brain regions, and edges, which denote statistical associations such as anatomical or effective connections. Graph edges can be categorized as weighted or unweighted, and either direct or indirect. Direct edges show a one-flow of information, where the activity of one node is influenced by the other. Indirect edges indicate two-way information flow between connected nodes. Strong and weak connections are distinguished by the line's weight, which indicates the edge's connectivity strength between two nodes. Thresholding removes weak connections.

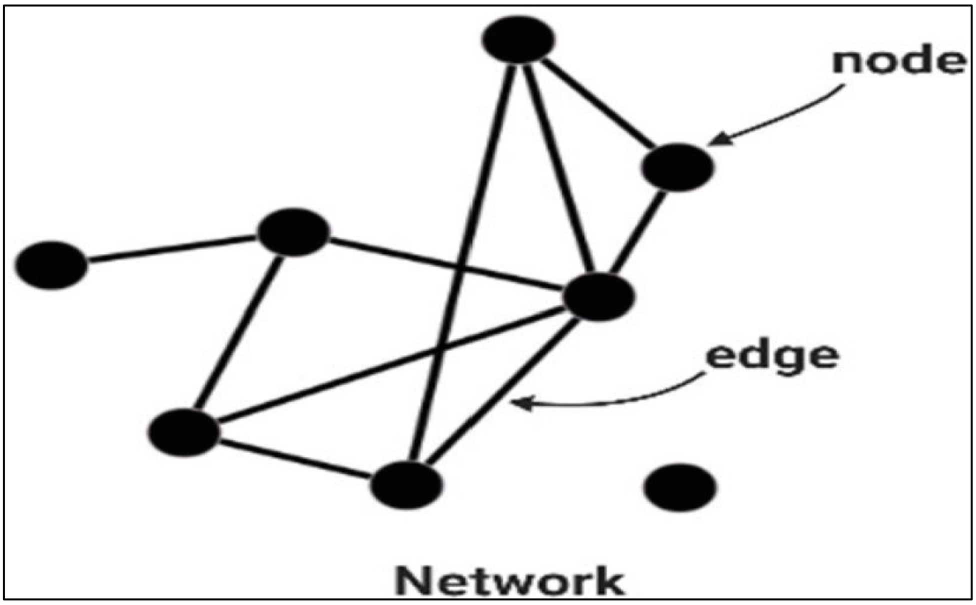

Figure 7 A representation of network consisting of nodes and edges.

In this study, implementing graph theory analysis based on Partial Directed Coherence (PDC) involves several steps, including the computation of PDC values from EEG data and the construction of a connectivity matrix. Once we have the connectivity matrix, we can apply graph theory to analyze the network properties and then calculate Partial Directed Coherence from EEG data to obtain a matrix representing the directional connectivity between two electrodes in a specific frequency band. Decide on a threshold value in order to create the binary from the connection matrix or weighted adjacency matrix. This step determines which connections are considered for further analysis. We use a graph library to construct a graph from the binary or weighted adjacency matrix. Each electrode becomes a node, and connections between electrodes become edges. Next step is to explain how we apply graph theoretical analysis apply in our study:

Once we compute the PDC matrices, which represent directional connections between EEG channels, these matrices are treated as adjacency matrices in graph theory. In this study we compute directionality the index, degree distribution and efficiency indexes:

3.10 Directionality Index

Each node in a graph corresponds to an EEG channel, and the degree represent the strength of connectivity between them, for these adjacency matrices. Before the directionality index is calculated, the node degree is calculated. The quantity of linkages between a node *i* to other nodes is known as its degree. This number is derived from adjacency matrix created using PDC values, which show which way information flows in a directed network. There are two categories for a node's degree: *in*-degree and *out*-degree. The *in*-degree (k_i^{in}) represents the quantity of entering connections, while the *out*-degree (k_i^{out}) denotes the quantity of outbound connections. These degrees can be expressed mathematically to quantify the flow of information [42]:

$$k_i^{in}(t) = \sum\nolimits_{j \in N_c} A_{ji}\,(t) \tag{7}$$

$$k_i^{out}(t) = \sum\nolimits_{j \in N_c} A_{ij}\,(t) \tag{8}$$

Where A_{ij} which denotes the adjacency matrix entry, need not always equal to A_{ji}. The ability of region to affects other is indicated by a node with a high *out*-degree value. Therefore, a node with a high *in-degree* value indicates that there might be a relationship between one region and another.

By calculating the difference between the out-degree and in-degree vectors, the directionality index (DI), which represents the information flow direction for each frequency band is calculated.

$$DI_i(t) = \sum k_i^{out}(t) - \sum k_i^{in}(t) \tag{9}$$

A feature matrix of size $N_T\ x\ N_c$, where N_T represents the number of time points and N_C represents the number of channels, was produced by storing the directionality index of the connectivity network at each time point in a row vector V_t: $1\ x\ N_c$.

3.10.1 Degree Distribution

The distribution of degree *P(k)* is a function indicates the percentage of nodes in the network with a degree *k.* In a network where nodes *i has* degree and k_i and there are nodes N.

$$P(k) = \frac{1}{N}\sum_{i=1}^{N} \delta(k - k_i) \qquad (10)$$

The Dirac delta function, denoted as **δ**(.) is 1 if $k=k_i$ and otherwise 0.

This gives a broad view of the network's topology and helps identify network types and node roles.

3.10.2 Efficiency Indexes

- Global Efficiency (E_{global}): Represents how effectively information is transferred over the whole network. It is calculated as:

$$E_{global} = \frac{1}{N(N-1)} \sum_{i \neq j} \frac{1}{d_{ij}} \qquad (11)$$

- Local Efficiency (E_{local} (i)): Reflects the efficiency of information transfer within the local neighborhood of node i. It is calculated as:

$$E_{local}(i) = \frac{1}{k_i(k_i - 1)} \sum_{j \neq k \in N(i)} \frac{1}{d_{jk}} \qquad (12)$$

Where $N(i)$ is the set neighbors of node i and $d_{jk}$is the shortest path length between nodes *j* and *k*.

From this study, they reveal how well information or resources flow through the network, both globally and locally. Efficiency metrics are crucial for understanding how stress impacts the functional organization of the brain. High or low values indicate by which measure stress affects overall and local information processing capability.

3.11 Microstate Analysis

The field of microstate analysis was first developed by Lehmann et al. [43] in 1987.They create discrete states out of the alpha frequency range(8-12Hz). Two surprising characteristics emerged when a multi-channel a state of rest to create a temporal sequence of scalp topographies, EEG data was transformed. Multi-channel recordings have more maps depicting the terrain of the scalp. These maps have a distinct 60-120ms temporal structure. A single topography gives way to another after a brief period of stability. These stable periods of a certain topography are called 'microstates. One method utilized in the field of electroencephalography (EEG) is microstate analysis is a technique to examine brief, steady patterns of brain activity. Brain electrical activity consisting of a series of brief, quasi-stable state known as microstate, which typically last for few milliseconds. The human brain operates with complex interaction between various regions, forming networks that change over time. The focus of EEG microstate analysis is the temporal evolution to potential topography. Topographical transitions and associated characteristics offer connection information.

The three steps of microstate analysis are parameter calculations, fitting, and clustering. The standard deviation of all channel potential at each site is equivalent to scalp potential strength is determined by measuring Global field power. Global field power (GFP) is computed for each electrode. To get average microstate topography, a modified K-mean clustering approach was used to topographies at GFP local minima with excellent signal-to-noise ratio were subjected to a modified version of K-mean clustering algorithm. Global field power [44] is computed as:

$$\mathrm{GFP}(t) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(V_i(t) - \overline{V}(t))^2} \tag{13}$$

where:

- N is the number of electrodes
- $V_i(t)$ represents the potential at time t at electrode i.
- $\overline{V}(t)$ is the average potential at time *t* across all electrodes.

This computes the root mean square deviation of the potential of all electrodes from the mean potential, providing a measure of how much the EEG field varies across the scalp at a given time.

From microstate data: the following parameters are extracted: transition probability, occurrence, temporal coverage, and duration.

The average microstate class duration, expressed in milliseconds, is called duration and that is:

$$D_i = \frac{\sum_{k=1}^{n_i} t_k}{n_i} \tag{14}$$

Here, n_i is the quantity of times the microstate i occurs and t_k is the duration of the k-th occurrence of microstate i

Number of occurrences per second is used to calculate the occurrence rate and refers to how frequently a particular microstate appears. The Occurrence O_i of a microstate i is simply the count of the number of times microstate observed:

$$O_i = n_i \tag{15}$$

The percentage of the entire analysis time that is devoted to time coverage is the proportion of total analysis time occupied by a particular set of microstate labels and computed as:

$$TC_i = \frac{\sum_{k=1}^{n_i} t_k}{T} \tag{16}$$

Transition probability refers to the likelihood of transitioning from one microstate to another during the course of EEG signal. We calculated the probabilities of transition from one microstate to another by:

$$P_{ij} = \frac{N_{ij}}{\sum_k N_{ik}} \tag{17}$$

Where, P_{ij} is the likelihood that microstate i will change to microstate *j*. N_{ij} is the quantity of times that microstate i changes into microstate j and $\sum_k N_{ik}$ is the total quantity of changes from the microstate i to another state *k*.

In this study, microstate analysis provides understanding of the dynamic temporal patterns of brain activity, while graph theoretical measures give an understanding of how these patterns manifest as network properties during stress and normal conditions.

3.12 Classification of Stress

The step after partial directed coherence is classification in the proposed methodology. The main contribution after feature extraction, labeled dataset of EEG features are classified into stress and normal. Features are used as inputs of classifiers. The classification of stress from EEG data is an important step in understanding how the brain reacts to stressful situation. In this study, three machine learnings algorithms are employed to classify EEG signals into normal and stress categories. Among these classifiers are k-Nearest Neighbors(k-NN), support vector machine (SVM), and Random Forest (RF).

3.12.1 Support Vector Machines (SVM)

Support Vector Machines (SVM) is an algorithm for machine learning that has recently been recognized as an effective categorization technique. SVM separates two classes with a function obtained from valuable training data. The objective is to generate classifiers using maximal input vectors that divide two regions that are SVM hyperplane functions. SVM is not restricted to distinguishing between two types of objects, there are various ways to create decision boundaries that divide the items into two groups. This method's objective is to find the best classifier function that can differentiate between two distinct data sets [45]. The separation function applied in this instance is linear.

$$g(x)=sign(f(x) \tag{18}$$

with $f(x) = \boldsymbol{w}^t\boldsymbol{x} + b, \boldsymbol{w}, \boldsymbol{x}, \in \boldsymbol{R}^n$ *and* $b \in \boldsymbol{R}$**,** the two parameters are w and b for which the goal is to find. The optimal hyperplane is positioned centrally between two object collections belonging to different types. Maximizing the margin or distance is the same as finding the optimal hyperplane separating the two groups of items from two classification. Samples positioned on the hyperplane are referred to as support vectors. This identification aims to identify the optimal hyperplane or classifier function from a set of functions.

For a linearly separable model, the goal is to find the hyperplane that optimizes the difference between two groups of data. In Fig., **w** represents the normal vector. The equation $\boldsymbol{w.x} + b=0$ defines the maximum margin hyperplane, while $\boldsymbol{w{\cdot}x}+b=\pm 1$ represents the parallel boundary hyperplanes. The distance between these boundary hyperplanes, also known as the margin is $\frac{2}{||w||}$. Therefore, finding the hyperplane with the maximum margin is equivalent to minimizing ||w||.

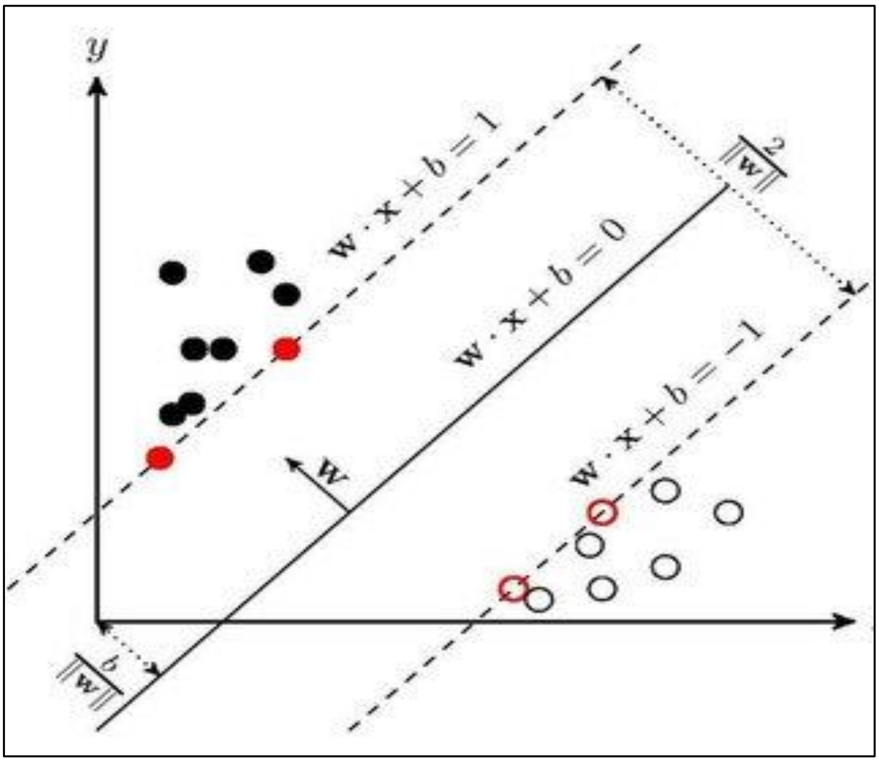

Figure 8 Linear Separable Support Vector Machine

3.12.2 Random Forests (RF)

An ensemble technique called Random Forests (RF) blends several decision trees. It enhances the effectiveness of one tree classifier by utilizing bootstrap aggregating method with randomization in the selection of data nodes during the classification process. The feature space is divided into *M* regions *Rm, 1≤ m ≤ M.* by a decision tree with M leaves. The prediction function for every tree is defined as:

$$f(x) = \sum_{m=1}^{M} c_m \Pi(x, R_m) \tag{19}$$

In the feature space, where *M* is the number of regions, R_m is a location suitable for m; c_m is a constant that fits with *m*:

$$\Pi(x, R_m) = \begin{cases} 1, & if\ x \in R_m \\ 0, & otherwise \end{cases} \tag{20}$$

The majority vote among all the trees determines the final categorization. An ensemble technique called Random Forest constructs numerous decision tress using different subsets of training data and features. It combines output of these trees to make more robust prediction. It can handle non-linear relationships, noisy data and high dimensional data like in our case EEG microstates.

3.12.3 K-Nearest Neighbor (KNN)

K-Nearest Neighbor operates on the principle that instances of a specific class are usually clustered with other instances of the identical class. Given a set of the feature space with labeled training examples and a scalar *k*, labeling an instance that is not labeled allows it to be classed and appears most frequently among the *k* nearest training samples. While various measures can be used to determine the space between occurrences, and the most widely utilized distance for this purpose is the Euclidean distance. Euclidean distance, which is represented by the following equation, is the kind of distance metric utilized in this technique:

$$L(x_i, x_j) = (\sum\nolimits_{i,j=1}^{n} ((|x_i - x_j|))^2)^{\frac{1}{2}} X \in R^n \tag{21}$$

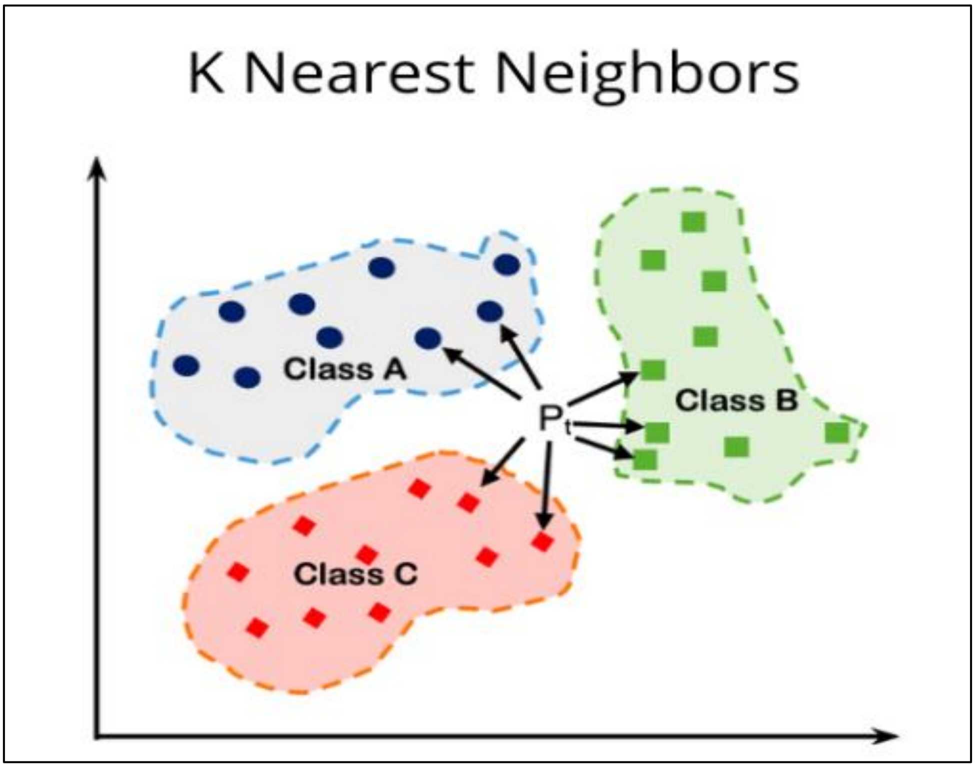

Figure 9 Representation of k-nearest neighbor

The above figure shows the representation of k-nearest neighbor, where k represents the nearest neighbor when classifying the data. It is an integer that is positive and usually small.

In this study, three classifiers of machine learning algorithms are used for analysis to take out the EEG task's properties. These classifiers are SVM, RF and KNN. The extracted features are split into 80-20 ratio for testing and training. An independent subject test with each classifier for 5-fold cross-validation is performed and also evaluated with the four-measure matrix. These include accuracy, precision, sensitivity and F-measure [46]. The formula displays the mathematical expression for each prediction. The outcomes of the confusion matrix are as follows:

- True Positive (TP): Numbers of labels accurately recognized as stress.
- True Negative (TN): Number of labels accurately recognized as normal
- False Positive (FP): Numbers of labels accurately recognized as stress
- False Negative (FN): Number of labels accurately recognized as rest

Below, formulas of accuracy, precision, recall and F-measure are discussed:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{22}$$

$$Precision = \frac{TP}{TP+FP} \tag{23}$$

$$Recall = \frac{TP}{TP+FN} \tag{24}$$

$$F\text{-}measure = 2\,\frac{Precision*Recall}{Precision+Recall} \tag{25}$$

Accuracy refers to the proportion of accurate predictions made across the complete dataset in two issues: stress and normal. Precision shows the correct measurement of positive prediction. Meanwhile, Recall relates to the complete measure of a classifier, that counts the actual stressors that are predicted in the dataset. The F-measure is a metric that combines both precision and recall to assess the effectiveness of detection results.

Chapter 4: Results

In this chapter, results are discussed that's we get after pre-processing, effective connectivity, graph theoretical measurement and after classification. The analysis aims to identifying the difference between stress and normal conditions based on brain network properties.

First, we summarize the outcomes of our preprocessing step:

4.1 EEG Data Pre-Processing

EEG data is loaded in the graphical user interface (GUI) of MATLAB i.e. EEGLAB in the file format. cnt, which is commonly used for EEG recordings. When we import the file, EEG structure contains all the information including the raw data, metadata like the channel information and sampling rate. In the preprocessing pipeline, we remove the noise. After removing the noise next step is to apply independent component analysis (ICA), where we remove the artifacts that are not relate with brain activity. This is important because by removing the non- brain related components that include heart, muscle, line noise and others information. We only want to focus on data that is used in our work and from which we get brain related information.

After loading the cnt files, first step is to select the desired channels for data-filtering. For connectivity analysis causal filter is used so the causal band-pass filter is applied to the dataset. It is used to preserve the temporal relationships in the data, which is important for effective connectivity analysis. Filtering is crucial step in preprocessing, as it helps to remove the noise and artifacts outside the frequency band of interest. In this study fig. 10 the signals in red color shows the noise where filtering is applied and frequency band is from 0 Hz to 30 Hz.

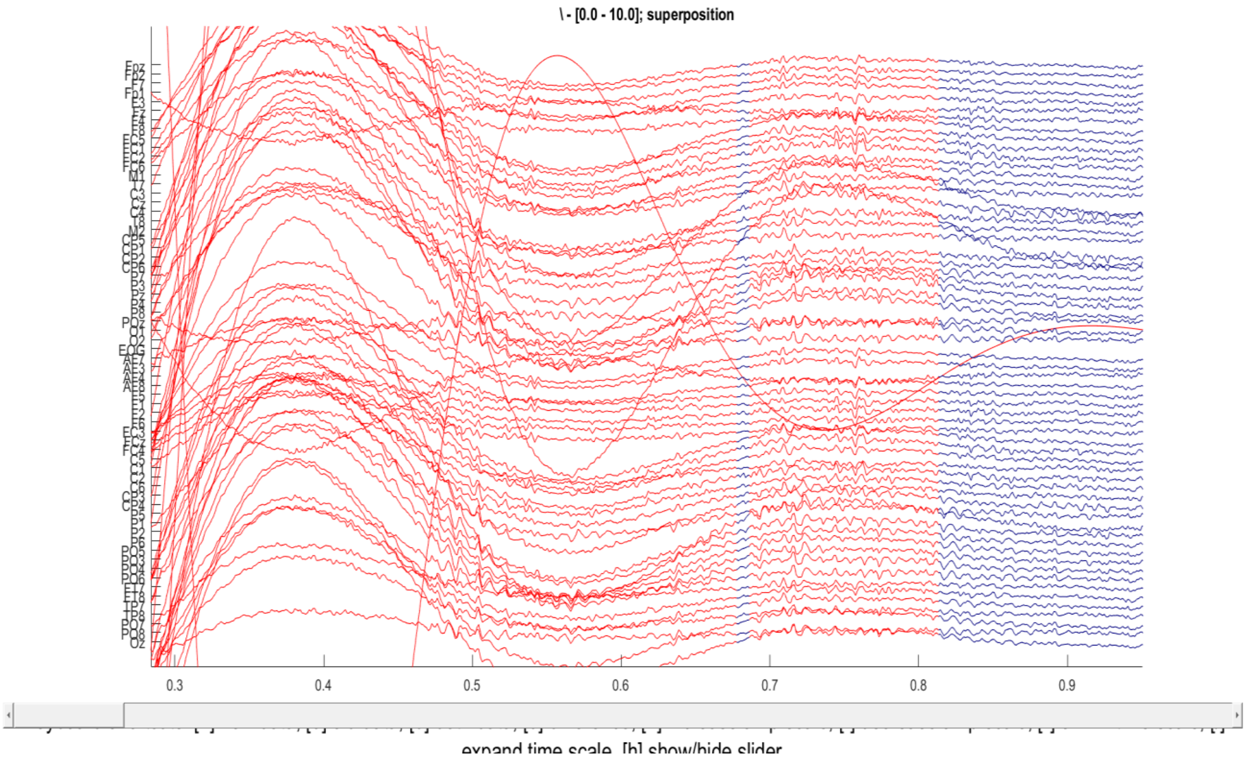

Figure 10 Noisy Signals

The low and high cut-off frequency for fir causal filter is set to remove low-frequency drifts and high- frequency drifts. The artifacts that are not in use, and that only cause noise are removed in fig.11. Moving towards further process, it's important to clean the data by removing any segment that is contaminated with noise or artifacts. In this study, when function identifies a segment where the voltage exceeds ±100µV and rejects from the dataset. After noise removal, the next step is to apply independent component analysis to identify and eliminate the data's artifacts that are not related to brain activity like muscle activity. All electrodes are reference to CPz and grounded AFz at so there is no need to re-reference the data. The sampling rate for this EEG data is 512Hz. Down sampling is the option for frequencies which are sampled at a frequency higher than 1000Hz. This improves the computational complexity, but when we perform it there are more chances of overlapping frequencies. Down sampling reduces the number of data points which can make it easier to perform subsequent analyses.

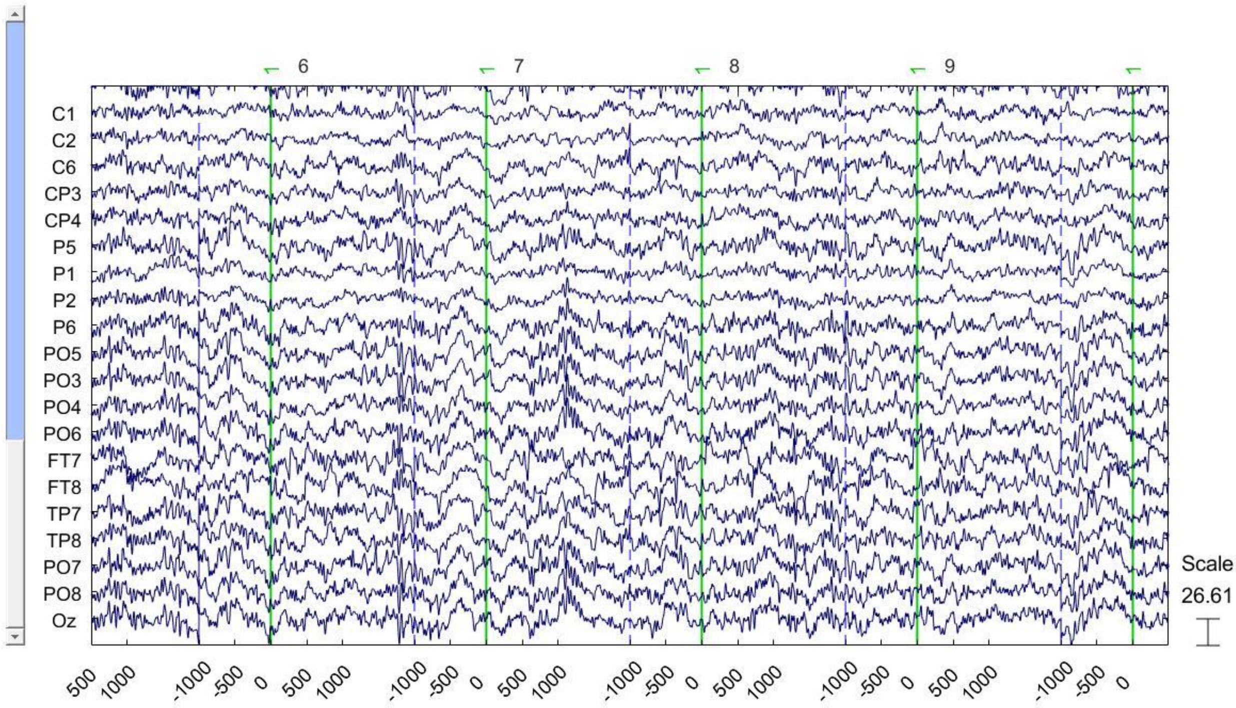

Figure 11 pre-process data

In Figure 11. Artifacts are shown that are removed after applying ICA due to their non-brain related information.

4.2 Effective Connectivity Results

Effective connectivity focuses on which brain region influences to other regions and gives insights into causal interaction between different brain regions. We employ partial directed coherence (PDC) in this work, as a connectivity measure to check which channel provides information about directional interaction between different brain region.

After preprocessing, a sliding window approach is used to break the continuous EEG data into smaller segments. Each segment contains a subset of EEG data that is analyzed for effective connectivity.

From the segment, the autoregressive model is fitted to the data to estimate the interaction between different channels. This model helps to find the dependencies of the current value of one channel to the past value of another channel. The coefficients from AR models are used to describe the relationship between the channels, which are used for computing the effective measure as in this study we mentioned PDC.

After the coefficients, PDC is computed for each window of that data. PDC tells about the directionality and strength of the connectivity between different EEG channels. It reveals how channel influence each other and in which direction. By PDC values, that is stored in the form of adjacency matrix, where elements represent the influence from one channel to another. And the values, which are computed is weights. These weights change when the sliding window is moved, so they reflect how brain connectivity evolves.

PDC is applied on 64 channels where they show connection between different brain regions. Weights are also shown and they represent the strength of how a channel is connected with another channel. When we perform pdc over time we compute adjacency matrices. From matrices of thirty-eight subjects file, we compute which channel is showing the maximum connection in their respective results. These files are already classified into stress and normal. After getting results on the basis of adjacency matrix over time we can compute which channel is connected with other channels in the network. In fig. 11 connectivity between different EEG channels is shown with weights and in this figure, the threshold is not applied they show connection among all channel.

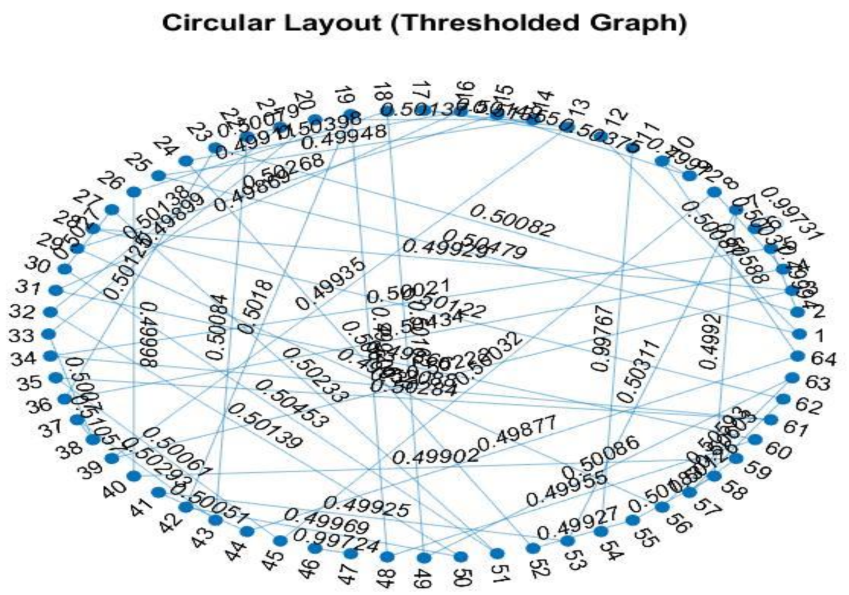

Figure 12 Connectivity Network with weights

In these results, effectivity connectivity analysis with PDC reveals which channel influence another channel in network. They show connection with all channel and threshold is not applied on these connections.

4.3 Graph Theoretical Analysis Outcomes

From effective connectivity where we perform partial directed coherence, we construct an adjacency matrix. So, the matrix represents the strength of the connection between different EEG channels, where the connection strength between two channels is represented by each entry. By analyzing degree centrality, global efficiency and local efficiency which is earlier discussed in chapter 3 of methodology where we discuss all these steps in graph theoretical analysis.

Fig. 12, shows the result of the graph theoretical analysis. In this result channels are differentiated on the basis of color that channel is the sink and which is the source. The channel which is shown in the blue is source and the channel that is shown in red is the sink. In fig. 11 all connected channels in the network are shown with weights. But in Fig. 12 threshold is applied to the weights that have a value above 0.90 are shown with their connection. Some of the channels are shown in red and blue dots but these are not connected with another channel in the network because the value of their weights is less than the threshold value. The channel which is shown in gray color dots are neutrals that are either sink or source.

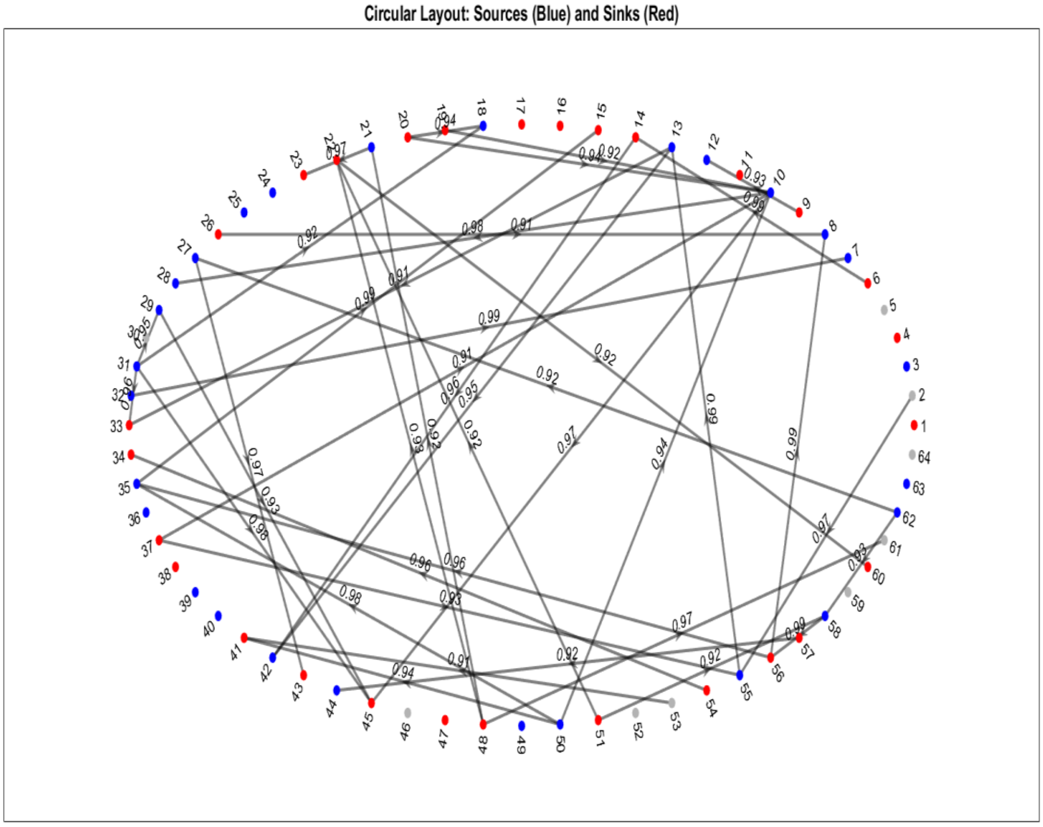

Figure 13 Graph theoretical analysis channel connection results

The above figure shown the results of how channels are connected but, in our study, when we classify stress and normal. In PDC we find out how channels show strength and influence between EEG signals. In fig. 13 results of the methods which are mentioned in the graph theoretical analysis section is shown. In this result we can estimate that in normal and stress which channel give high value. In degree centrality where we can tell that which cannel (node) is connected with other channel in this case normal and stress is differentiable with respect to the blue and red bar color bar graph. Normal channel Pz gives a high degree centrality value that reflects cognitive or normal state condition. But in the case of stress channel PO5 shows a high degree of centrality, it plays a significant role in receiving and sending a lot of signals.

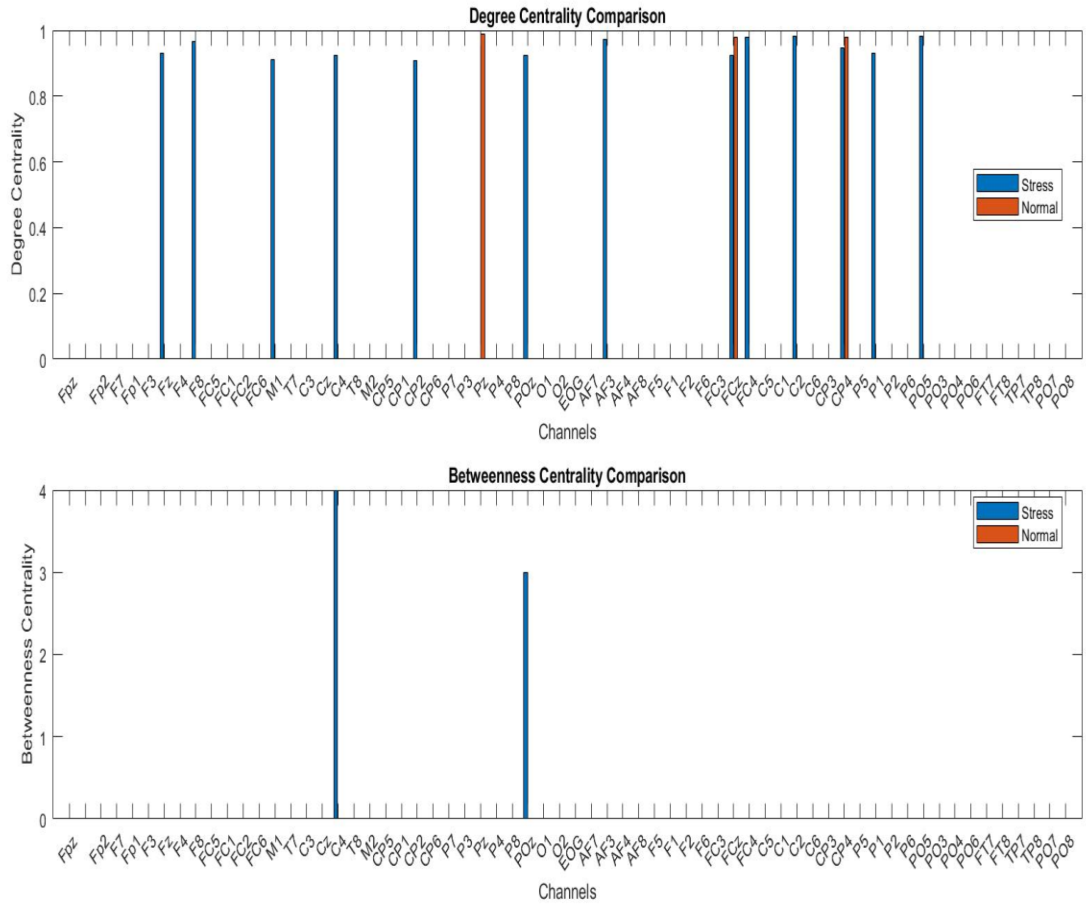

Figure 14 Graph Theory degree and betweenness result

As in the above process, we apply PDC and graph theoretical analysis where PDC is applied to find out the strength and influence of channels and it reveals pairwise connections. On the other hand, graph theoretical analysis provides quantitative measurements that summarize the network.

4.4 Microstate Analysis

Microstate analysis is a technique for determining the stable patterns of brain activity within a continuous EEG recording. This study's goal is to apply after-graph theoretical analysis is identify stable and recurring patterns. Because this approach is helpful in understanding brain activity dynamics. The k-means clustering method is used to identify the microstate. These topographic maps represent that how state change from one to another state. Results of this study for microstate analysis show with respect to normal and stress shown in Fig 14.

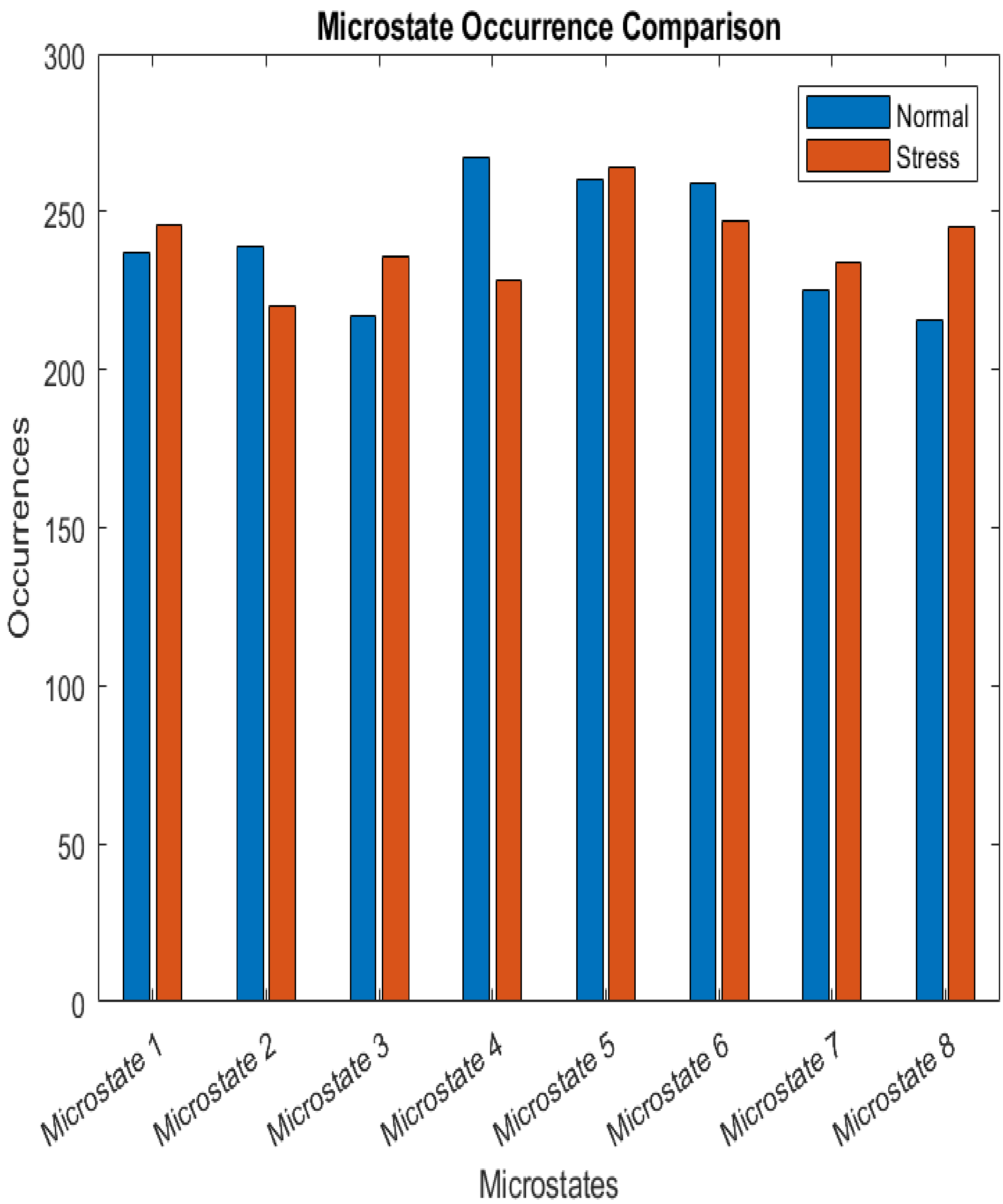

Figure 15 Occurrence of stress and normal

Fig. 14, occurrence for normal and stress in microstate is shown. After performing graph theoretical analysis, where metrics like degree of centrality and clustering coefficient values are used to understand the brain network properties during each microstate. These 8 microstates are determined by clustering EEG data into distinct spatial configurations using k-means clustering. In this study 8 microstates are shown which corresponds to a unique configuration of brain activity and the states represent the recurrent patterns that the brain exhibits during EEG recording. In Fig. 14 microstate 4 shows the highest occurrence among other states for normal conditions this state is linked with baseline cognitive processing. From graph theoretical measures it shows that brain network configuration is more stable for non-stressed in this state.

For stress conditions microstate 5 is shown with a high occurrence value suggesting that this brain configuration is more prominent under stress conditions. Microstate 5 represents that brain activity is activated during stress response related to emotional processing.

In fig. 15 duration for normal and stress is shown. Duration is linked with the length of the time brain remains in a particular microstate before transitioning to another state. Microstate 4 shows a high duration for normal it suggesting that this certain state is a more relaxed, low demand cognitive state. On the other hand, microstate 5 shows a high duration value that reflects the activation of neural circuits involved in stress response.

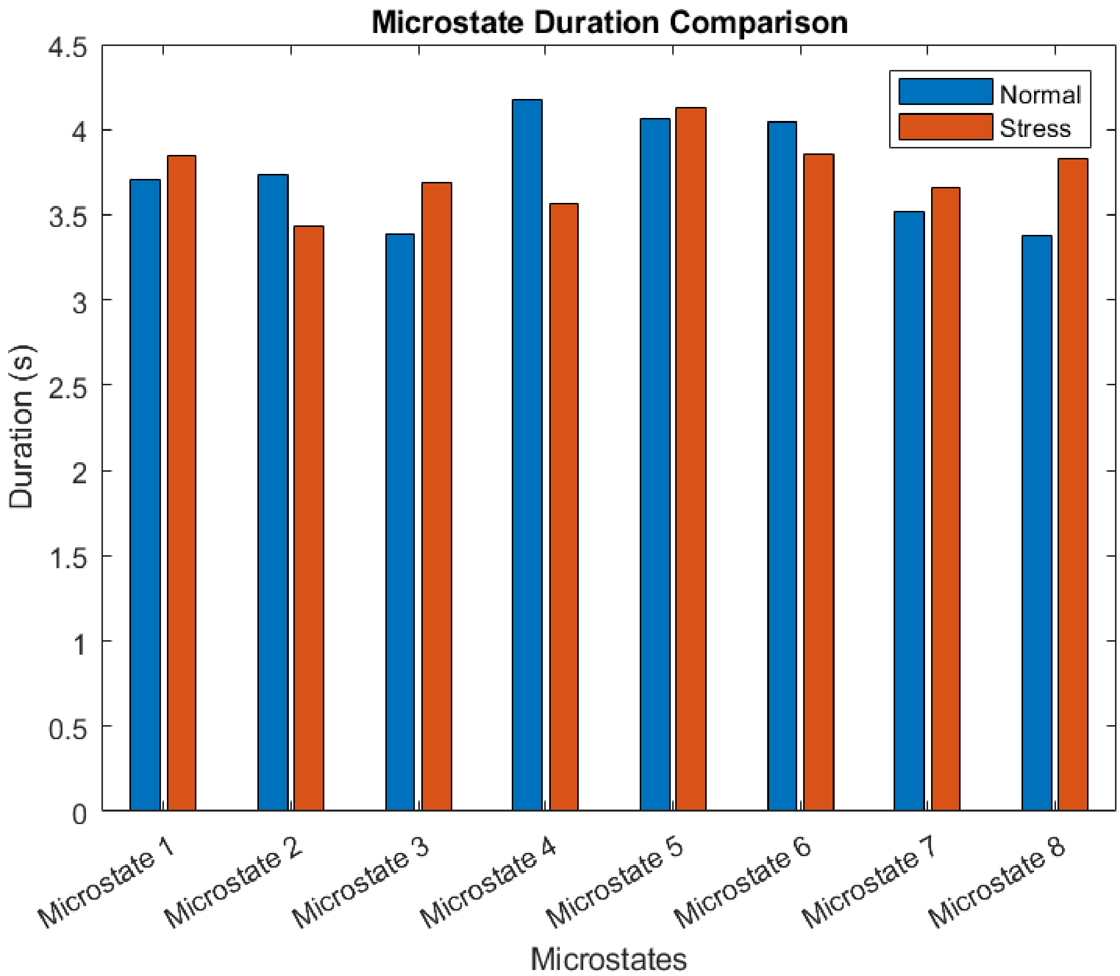

Figure 16 Duration of stress and normal

Fig. 15 is about the state transition of stress and normal. State transition refers to how often the brain shifts from microstate from one state to another. High transition linked with a particular microstate in the brain frequently moves from one state to another state. This shows how dynamic and reactive the brain is during different conditions.

In normal conditions, microstate 4 shows frequent transition, which suggests that the brain is actively switching to this state during period of low cognitive demand. For stress microstate 5 shows frequent transition that suggest the brain is frequently engaging in a state that is related with stress-related processing.

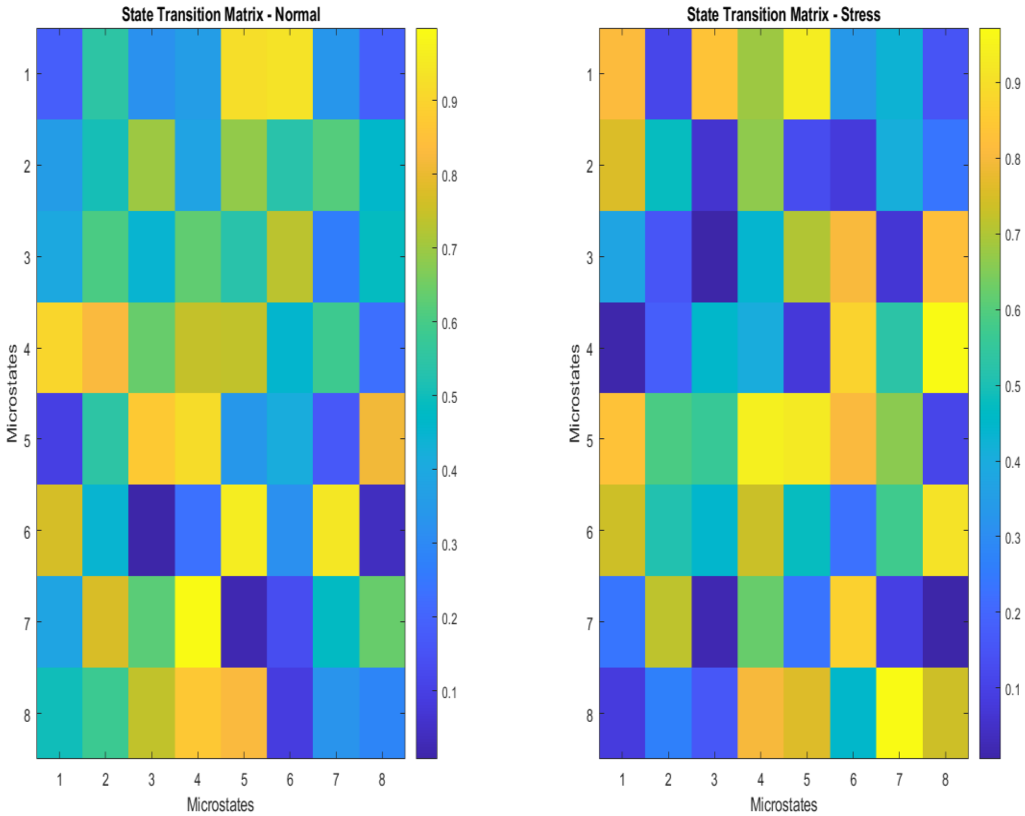

Figure 17 State Transition of stress and normal

Figure 18. is about topographies map of normal and stress. This shows the 8 states of normal and stress and by these we can estimate that which state is highly occurring and change their transition in milliseconds.

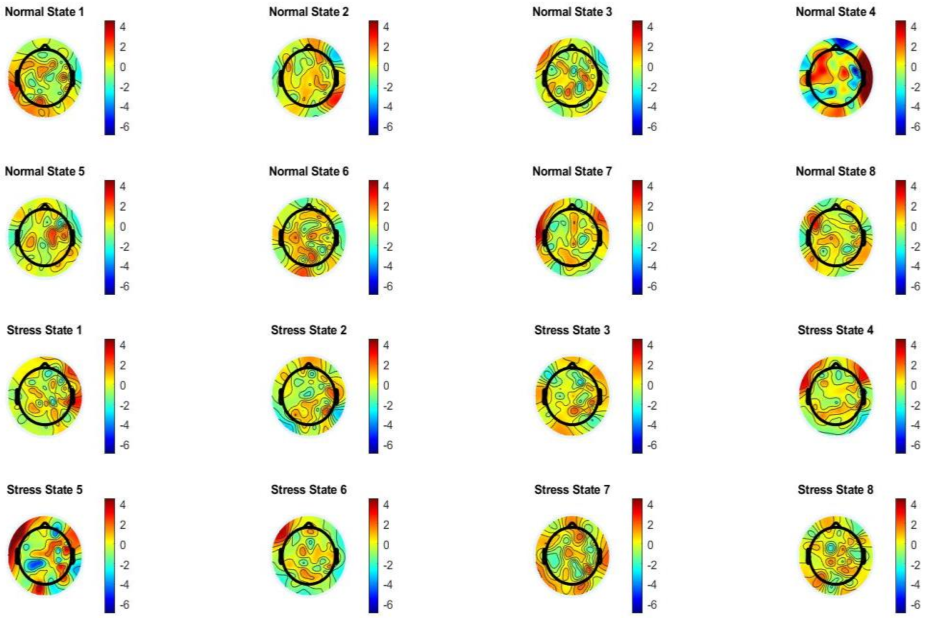

Figure 18 Topographies Map of Microstates

Figure 19. shows the states with high duration and occurrence maps for stress and normal. In normal state 4 shows high occurrence and in stress microstate 5 shows high occurrence.

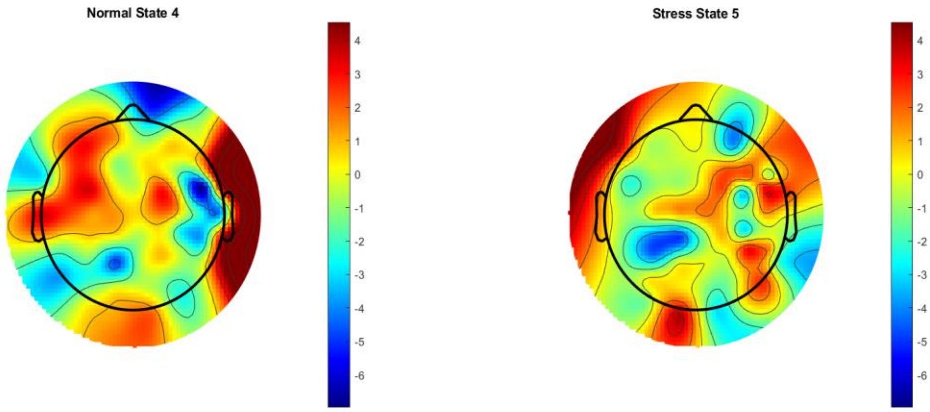

Figure 19 Maps with high occurrence states

4.5 Classification Accuracy and Performance Metrics

Random forest shows high accuracy because it is a more adaptable and powerful model to manage the complex, non-linear nature of EEG microstate data when the variability is present in normal and stress states.

KNN performs better than SVM because EEG microstate from reasonably distinct clusters, but KNN distance- based still struggle with high-dimensionality and noise. SVM is good for linear separations, but less suited for complex, noisy and non-linearly separable data like EEG. Random Forest outperforms because of its robustness, and capacity to model non-linear relationships, while KNN performs reasonably well but struggles with dimensionality. SVM being more sensitive to noise, shows the lowest accuracy in the EEG microstate classification task.

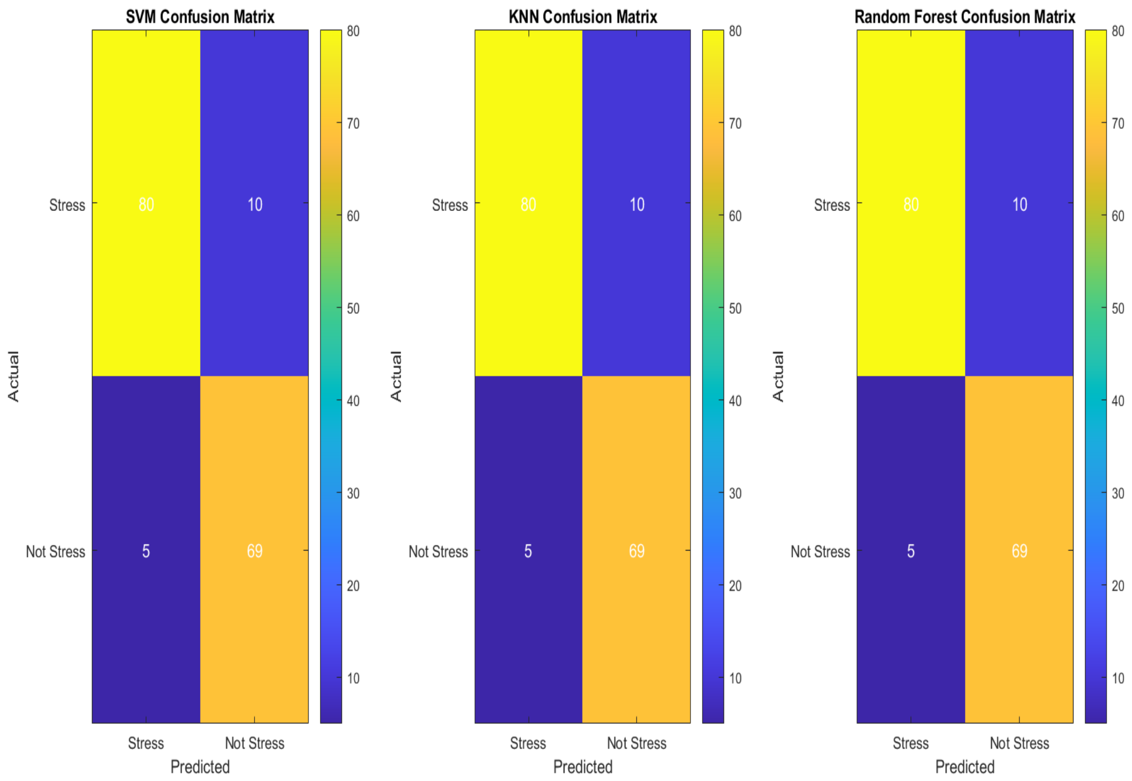

Figure 20 Confusion matrix of stress and normal

Fig 19. Shows these results show the recall, precision and F1_score of the mentioned classifiers.

	Classifier	**Accuracy**	**Recall**	**Precision**	**F1_Score**
1	**SVM**	84	0.85	0.85	0.85
2	**KNN**	85	0.78	0.82	0.79
3	**RF**	96	0.96	0.97	0.96

Figure 21 Classification Results

Chapter 5: Discussion

This chapter is about the implementation of the proposed design and future tasks. We first discuss our methodology where pre-processing is used to remove the noise and for this filtering is applied. After noise, artifacts are removed by independent component analysis. These artifacts include muscle and heart activity and line noise that affects the EEG signals. By removing these artifacts partial directed coherence, utilized in order to estimate the strength of the channel and directionality among network. By PDC measures they allow us to construct effective brain networks in time series. Connectivity networks are constructed by adjacency matrix which are computed by PDC. These connectivity networks provide insights into channels connection in the network. Next is the graph theoretical analysis that quantifies these networks and alters the degree and betweenness. Graph theory helps in translate the connectivity information from PDC into matrices and that describe the structure of the brain network. Degree and betweenness is computed where degree measure the number of connections a channel has in the network and betweenness quantifies the shortest path with other channels. Channels with higher betweenness reveal that this is the bridge between networks for communicating with other channels. In this study channel 19 shows high degree that act as an influential region in term of connectivity and channel 23 show high betweenness that indicates as a mediator between different brain regions. It plays an essential role in facilitating a communication between different channels in the network that are not directly related to each other. PDC and graph theory talks about connectivity patterns between brain region and on the other hand, microstate analysis is used to give a dynamic snapshot of brain activity, capturing brain states dynamically over time. In microstate analysis, occurrence and duration is the parameters used to figure out that this state occurs for stress and normal categories and we then perform classification using classifiers like K-Nearest Neighbor (KNN), Support Vector Machine (SVM), and Random Forest (RF). Among all RF shows high accuracy differentiating stress and normal. In future work using effective connectivity measures helps in directionality and shows the strength of channels among networks. Using PDC with graph theory and microstate analysis helps in finding the region of stress but advanced classifiers used as CNN and deep neural networks are used to capture complex patterns in microstate transitions. By using deep learning algorithms where we trained a model for stress detection and this can be used in real-time. In real-time we monitor brain activity that induces stress and assess it so we can get the root cause of stress.

Chapter 6: Conclusion

A comprehensive approach is applied to EEG workflow that integrates the analysis of the independent component; this is the step where we perform pre-processing before the independent component analysis which is applied to raw EEG data to remove the unwanted frequency component using band-pass filters. Down sample the data from 512 to 64 to make the data more manageable and reduce computational load. After down sampling, channel selection is another task but in the initial steps work with 64 channels. ICA is applied to remove the noise sources and separate the neural signals from non-neural artifacts. After the cleaning process, EEG data is used to apply partial directed coherence to quantify the direction of information flow between brain regions and how one region influences the other, these directions are shown by constructing the effective connectivity networks where the adjacency network shows which channel influences the other. These networks were then quantified using graph theoretical analysis to compute key metrices such as degree centrality and betweenness. In graph theoretical analysis, from channel direction, we can compute which channel is connected with another channel in the network. Their connection is shown with weights and based on these the channels that have high weight values after applying thresholding we can compute these are the channels in the network that influence others. Other key points are degree and betweenness which measures how well the node(channels) is connected with others in the network. A high degree indicates that the node is central to the network and show their direct connections and betweenness is about a node lies on the shortest path between two other nodes. Graph theory is all about direction and interaction among networks. The next step is to pass that graph theoretical results to microstate analysis by which based on states we can figure out which relate to stress and normal. In this study, microstate 4 relates to the normal state, and the microstate 5 refers to the stress state, and these states are formed on the results we compute from graph theory. The next step is to classify stress and normal from the classifiers which are above mentioned. RF, kNN and SVM give the 94 %, 90% and 87% accuracy, respectivel

References

1. Goldstein, D. S., & Kopin, I. J. (2007). Evolution of concepts of stress. *Stress*, *10*(2),109-120.https://doi.org/10.1080/10253890701288935
2. Schneiderman N, Ironson G, Siegel SD. Stress and health: psychological, behavioral, and biological determinants. Annu Rev Clin Psychol. 2005; 1:607-28. doi: 10.1146/annurev.clinpsy.1.102803.144141. PMID: 17716101; PMCID: PMC2568977.
3. Córdova A, Caballero-García A, Drobnic F, Roche E, Noriega DC. Influence of Stress and Emotions in the Learning Process: The Example of COVID-19 on University Students: A Narrative Review. Healthcare (Basel). 2023 Jun 17;11(12):1787. doi: 10.3390/healthcare11121787. PMID: 37372905; PMCID: PMC10298416.
4. Attia M, Ibrahim FA, Elsady MA, Khorkhash MK, Rizk MA, Shah J, Amer SA. Cognitive, emotional, physical, and behavioral stress-related symptoms and coping strategies among university students during the third wave of COVID-19 pandemic. Front Psychiatry. 2022 Sep 16; 13:933981. doi: 10.3389/fpsyt.2022.933981. PMID: 36186888; PMCID: PMC9523087.
5. Dimsdale JE. Psychological stress and cardiovascular disease. J Am Coll Cardiol. 2008 Apr 1;51(13):1237-46. doi: 10.1016/j.jacc.2007.12.024. PMID: 18371552; PMCID: PMC2633295.
6. Jharna Agrawal *et al* 2021 *IOP Conf. Ser.: Mater. Sci. Eng.* **1116** 012134**DOI** 10.1088/1757-899X/1116/1/012134
7. Tchoe Y, Bourhis AM, Cleary DR, Stedelin B, Lee J, Tonsfeldt KJ, Brown EC, Siler DA, Paulk AC, Yang JC, Oh H, Ro YG, Lee K, Russman SM, Ganji M, Galton I, Ben-Haim S, Raslan AM, Dayeh SA. Human brain mapping with multithousand-channel PtNRGrids resolves spatiotemporal dynamics. Sci Transl Med. 2022 Jan 19;14(628):eabj1441. doi: 10.1126/scitranslmed.abj1441. Epub 2022 Jan 19. PMID: 35044788; PMCID: PMC9650779.
8. Aderinto, N., Olatunji, G., Muili, A. *et al.* A narrative review of non-invasive brain stimulation techniques in neuropsychiatric disorders: current applications and future directions. *Egypt J Neurol Psychiatry Neurosurg* **60**, 50 (2024). https://doi.org/10.1186/s41983-024-00824-w
9. Šimić G, Tkalčić M, Vukić V, Mulc D, Španić E, Šagud M, Olucha-Bordonau FE, Vukšić M, R Hof P. Understanding Emotions: Origins and Roles of the Amygdala. Biomolecules. 2021 May 31;11(6):823. doi: 10.3390/biom11060823. PMID: 34072960; PMCID: PMC8228195.

10. Rakesh Ranjan, Bikash Chandra Sahana, Ashish Kumar Bhandari, Ocular artifact elimination from electroencephalography signals: A systematic review, Biocybernetics and Biomedical Engineering, Volume 41, Issue 3,2021, Pages 960-996, ISSN 0208-5216, https://doi.org/10.1016/j.bbe.2021.06.007.
11. Marta Cerina *et al* 2023 *Prog. Biomed. Eng.* **5** 032002**DOI** 10.1088/2516-1091/acce12
12. Tivadar, R. I., & Murray, M. M. (2019). A Primer on Electroencephalography and Event-Related Potentials for Organizational Neuroscience. Organizational Research Methods, 22(1), 69-94. https://doi.org/10.1177/1094428118804657
13. Vijean, V., Hariharan, M., Saidatul, A. and Yaacob, S., 2011, October. Mental tasks classifications using S-transform for BCI applications. In 2011 IEEE Conference on Sustainable Utilization and Development in Engineering and Technology (STUDENT) (pp. 69-73).IEEE.
14. Elzeiny, S. and Qaraqe, M., 2018, August. Blueprint to Workplace Stress Detection Approaches.In 2018 International Conference on Computer and Applications (ICCA) (pp. 407- 412).IEEE.
15. Horwitz TB, Balbona JV, Paulich KN, Keller MC. Evidence of correlations between human partners based on systematic reviews and meta-analyses of 22 traits and UK Biobank analysis of 133 traits. Nat Hum Behav. 2023 Sep;7(9):1568-1583. doi: 10.1038/s41562-023-01672-z. Epub 2023 Aug 31. PMID: 37653148; PMCID: PMC10967253.
16. Horwitz, B. The Elusive Concept of Brain Connectivity. *Neuroimage* **2003**, *19*, 466–470
17. Nunez, P.L.; Srinivasan, R. *Electric Fields of the Brain: The Neurophysics of EEG*; Oxford University Press: Oxford, MI, USA, 2006.
18. Friston, K.J. Functional and effective connectivity in neuroimaging: A synthesis. *Hum. Brain Mapp.* **1994**, *2*, 56–78.
19. Friston, Karl J. "Functional and effective connectivity: a review." *Brain connectivity* 1, no. 1 (2011): 13-36.
20. Xia, L.; Malik, A.S.; Subhani, A.R. A physiological signal-based method for early mental-stress detection. *Biomed. Signal Process. Control* **2018**, *46*, 18–32.
21. Al-Shargie, F. Early Detection of Mental Stress Using Advanced Neuroimaging and Artificial Intelligence. *arXiv* **2019**, arXiv:1903.08511.
22. Zanetti, M.; Faes, L.; De Cecco, M.; Fornaser, A.; Valente, M.; Guandalini, G.; Nollo, G. Assessment of mental stress through the analysis of physiological signals acquired from

wearable devices. In Proceedings of the Italian Forum of Ambient Assisted Living, Lecce, Italy, 2–4 July 2018; pp. 243–256.

23. Alonso, J.; Romero, S.; Ballester, M.; Antonijoan, R.; Mañanas, M. Stress assessment based on EEG univariate features and functional connectivity measures. *Physiol. Meas.* **2015**, *36*, 1351.
24. Al-Shargie, F.; Tariq, U.; Alex, M.; Mir, H.; Al-Nashash, H. Emotion recognition based on fusion of local cortical activations and dynamic functional networks connectivity: An EEG study. *IEEE Access* **2019**, *7*, 143550–143562.
25. Darzi, A.; Azami, H.; Khosrowabadi, R. Brain functional connectivity changes in long-term mental stress. *J. Neurodev. Cogn.* **2019**, *1*, 16–41.
26. Darzi, A.; Azami, H.; Khosrowabadi, R. Brain functional connectivity changes in long-term mental stress. *J. Neurodev. Cogn.*
27. Khosrowabadi, R.; Quek, C.; Ang, K.K.; Tung, S.W.; Heijnen, M. A Brain-Computer Interface for classifying EEG correlates of chronic mental stress. In Proceedings of the 2011 International Joint Conference on Neural Networks, San Jose, CA, USA, 31 July–5 August 2011; pp. 757–762.
28. Al-Shargie, F.; Tariq, U.; Alex, M.; Mir, H.; Al-Nashash, H. Emotion recognition based on fusion of local cortical activations and dynamic functional networks connectivity: An EEG study. *IEEE Access* **2019**, *7*, 143550–143562.
29. Khosrowabadi, R. Stress and Perception of Emotional Stimuli: Long-term Stress Rewiring the Brain. *Basic Clin. Neurosci.* **2018**, *9*, 107.
30. Al-Ezzi A, Kamel N, Al-Shargabi AA, Al-Shargie F, Al-Shargabi A, Yahya N, Al-Hiyali MI. Machine learning for the detection of social anxiety disorder using effective connectivity and graph theory measures. Front Psychiatry. 2023 May 9;14:1155812. doi: 10.3389/fpsyt.2023.1155812. Erratum in: Front Psychiatry. 2023 Jul 24;14:1257713. doi: 10.3389/fpsyt.2023.1257713. PMID: 37255678; PMCID: PMC10226190.
31. Xefteris, V.-R.; Tsanousa, A.; Georgakopoulou, N.; Diplaris, S.; Vrochidis, S.; Kompatsiaris, I. Graph Theoretical Analysis of EEG Functional Connectivity Patterns and Fusion with Physiological Signals for Emotion Recognition. *Sensors* **2022**, *22*, 8198. https://doi.org/10.3390/s22218198
32. Kim K, Duc NT, Choi M, Lee B. EEG microstate features for schizophrenia classification. PLoS One. 2021 May 14;16(5):e0251842. doi: 10.1371/journal.pone.0251842. PMID: 33989352; PMCID: PMC8121321.

33. Shaleen Bhatnagar, Sarika Khandelwal, Shruti Jain, Harsha Vyawahare,A deep learning approach for assessing stress levels in patients using electroencephalogram signals,Decision Analytics Journal,Volume 7,2023,100211,ISSN 2772 6622,https://doi.org/10.1016/j.dajour.2023.100211.
34. Christoph M. Michel, Thomas Koenig,EEG microstates as a tool for studying the temporal dynamics of whole-brain neuronal networks: A review,NeuroImage,Volume 180, Part B,2018,Pages 577-593,ISSN 1053-8119,https://doi.org/10.1016/j.neuroimage.2017.11.062.

35. Kheirati Roonizi, A., Sassi, R. ECG signal decomposition using Fourier analysis. *EURASIP J. Adv. Signal Process.* **2024**, 79 (2024). https://doi.org/10.1186/s13634-024-01171.

36. João D Semedo, Evren Gokcen, Christian K Machens, Adam Kohn, Byron M Yu, Statistical methods for dissecting interactions between brain areas, Current Opinion in Neurobiology, Volume 65,2020,Pages 59-69,ISSN 0959-4388 https://doi.org/10.1016/j.conb.2020.09.009.

37. Ahmadi H, Fatemizadeh E, Motie-Nasrabadi A. A Comparative Study of Correlation Methods in Functional Connectivity Analysis Using fMRI Data of Alzheimer's Patients. J Biomed Phys Eng. 2023 Apr 1;13(2):125-134. doi: 10.31661/jbpe.v0i0.2007-1134. PMID: 37082543; PMCID: PMC10111107.

38. Björn Schelter, Jens Timmer, Michael Eichler, Assessing the strength of directed influences among neural signals using renormalized partial directed coherence, Journal of Neuroscience Methods, Volume 179, Issue 1,2009,Pages 121-130,ISSN 0165-0270,https://doi.org/10.1016/j.jneumeth.2009.01.006.
39. Baccalá LA, Sameshima K. Partial Directed Coherence and the Vector Autoregressive Modelling Myth and a Caveat. Front Netw Physiol. 2022 Apr 28;2:845327. doi: 10.3389/fnetp.2022.845327. PMID: 36926097; PMCID: PMC10012995.
40. Colin R. Buchanan, Mark E. Bastin, Stuart J. Ritchie, David C. Liewald, James W. Madole, Elliot M. Tucker-Drob, Ian J. Deary, Simon R. Cox,The effect of network thresholding and weighting on structural brain networks in the UK Biobank,NeuroImage,Volume 211,2020, 116443,ISSN 1053-8119,https://doi.org/10.1016/j.neuroimage.2019.116443.

41. Jelfs B, Chan RHM. Directionality indices: Testing information transfer with surrogate correction. Phys Rev E. 2017 Nov;96(5-1):052220. doi: 10.1103/PhysRevE.96.052220. Epub 2017 Nov 27. PMID: 29347680.
42. Asha S.A, Sudalaimani C, Devanand P, Alexander G, Arya Maniyan Lathikakumari, Sanjeev V Thomas, Ramshekhar N Menon, Analysis of EEG microstates as biomarkers in neuropsychological processes – Review, Computers in Biology and Medicine, Volume 173,2024,108266, ISSN 0010-4825, https://doi.org/10.1016/j.compbiomed.2024.108266.
43. Schiller B, Sperl MFJ, Kleinert T, Nash K, Gianotti LRR. EEG Microstates in Social and Affective Neuroscience. Brain Topogr. 2024 Jul;37(4):479-495. doi: 10.1007/s10548-023-00987-4. Epub 2023 Jul 31. PMID: 37523005; PMCID: PMC11199304.
44. Skrandies, W. Global field power and topographic similarity. *Brain Topogr* **3**, 137–141 (1990). https://doi.org/10.1007/BF01128870
45. Andrew AM. An Introduction to support vector machines and other kernel-based learning methods. Kybernetes. 2001. https://doi.org/10.1108/k.2001.30.1.103.6.
46. Hag, A.; Handayani, D.; Pillai, T.; Mantoro, T.; Kit, M.H.; Al-Shargie, F. EEG Mental Stress Assessment Using Hybrid Multi-Domain Feature Sets of Functional Connectivity Network and Time-Frequency Features. *Sensors* **2021**, *21*, 6300. https://doi.org/10.3390/s21186300