UNIVERSITI TEKNOLOGI MARA

DETERMINATION AND CLASSIFICATION OF HUMAN STRESS INDEX USING NON-PARAMETRIC ANALYSIS OF EEG SIGNALS

NORIZAM BIN SULAIMAN

PhD

November 2015

UNIVERSITI TEKNOLOGI MARA

DETERMINATION AND CLASSIFICATION OF HUMAN STRESS INDEX USING NON-PARAMETRIC ANALYSIS OF EEG SIGNALS

NORIZAM BIN SULAIMAN

Thesis submitted in fulfillment
of the requirements for the degree of
Doctor of Philosophy

Faculty of Electrical Engineering

November 2015

CONFIRMATION BY PANEL OF EXAMINERS

I certify that a Panel of Examiners has met on 14th May 2015 to conduct the final examination of Norizam Sulaiman on his Doctor of Philosophy thesis entitled "Determination and Classification of Human Stress Index Using Non-Parametric Analysis of EEG Signals" in accordance with Universiti Teknologi MARA Act 1976 (Akta 173). The Panel of Examiners recommends that the student be awarded the relevant degree. The panel of Examiners was as follows:

Datin Shah Rizam Mohd Shah Baki, PhD
Professor
Faculty of Electrical Engineering, UiTM
(Chairman)

Nooritawati Md Tahir, PhD
Associate Professor
Faculty of Electrical Engineering, UiTM
(Internal Examiner)

Herlina Abdul Rahim, PhD
Associate Professor Ir
Faculty of Electrical Engineering, UTM
(External Examiner)

Deepak Ghodgaonkar, PhD
Associate Professor
Dhirubhai Ambani Institute of Information and Communication Technology, India
(External Examiner)

SITI HALIJJAH SHARIFF, PhD
Associate Professor
Dean
Institute of Graduates Studies
Universiti Teknologi MARA
Date: 3rd November, 2015

AUTHOR'S DECLARATION

I declare that the work in this thesis was carried out in accordance with the regulations of Universiti Teknologi MARA. It is original and is the result of my own work, unless otherwise indicated or acknowledged as referenced work. This thesis has not been submitted to any other academic institution or non-academic institution for any other degree or qualification.

I, hereby, acknowledge that I have been supplied with the Academic Rules and Regulations for Post Graduate, Universiti Teknologi MARA, regulating the conduct of my study and research.

Name of Student	:	Norizam Bin Sulaiman
Student ID No.	:	2008728647
Programme	:	EE990
Faculty	:	Electrical Engineering
Thesis Title	:	Determination and Classification of Human Stress Index using Non-Parametric Analysis of EEG Signals

Signature of Student	:	………………………………………………
Date	:	November 2015

ABSTRACT

Regardless of type of stress, either mental stress, emotional stress or physical stress, it definitely affects human lifestyle and work performance. There are two prominent methods in assessing stress which are psychological assessment (qualitative method) and physiological assessment (quantitative method). This research proposes a new stress index based on Electroencephalogram (EEG) signals and non-parametric analysis of the signals. In non-parametric method, the EEG features that might relate to stress are extracted in term of Asymmetry Ratio (AR), Relative Energy Ratio (RER), Spectral Centroids (SC) and Spectral Entropy (SE). The selected features are fed to the k-Nearest Neighbor (k-NN) classifier to identify the stressed group among the four experimental groups being tested. The classification results are based on accuracy, sensitivity and specificity. To support the classification results using k-NN classifier, the clustering techniques using Fuzzy C-Means (FCM) and Fuzzy K-Means (FKM) are implemented. To ensure the robustness of the classifier, the cross-validation technique using k-fold and leave-one-out is performed to the classifier. The assignment of the stress index is verified by applying Z-score technique to the selected EEG features. The experiments established a 3-level index (Index 1, Index 2 and Index 3) which represents the stress levels of low stress, moderate stress and high stress at overall classification accuracy of 88.89%, classification sensitivity of 86.67 % and classification specificity of 100%. The outcome of the research suggests that the stress level of human can be determined accurately by applying SC on the ratio of the Energy Spectral Density (ESD) of Beta and Alpha bands of the brain signals. The experimental results of this study also confirm that human stress level can be determined and classified precisely using physiological signal through the proposed stress index. The high accuracy, sensitivity and specificity of the classifier might also indicate the robustness of the proposed method.

ACKNOWLEDGEMENT

First of all, I would like to thank ALLAH Almighty, Lord of the Universe, for giving me a good health and strength to come out with this manuscript. Secondly, the great appreciation goes to my beloved supervisor, Professor Dr. Mohd Nasir Taib, who are always ready to guide me to correct my path and direction in my challenging PhD journey. His valuable ideas, great patience, high tolerance and good leadership skills provide me a high spirit to face any difficulty during my research work. Besides, I had a lot of fruitful discussions with him.

Also counted in this acknowledgements are my research partners and members under Advanced Signal Processing Research Group, Faculty of Electrical Engineering, Universiti Teknologi MARA such as Associate Professor Zunairah Hj Murat, Sahrim Lias, Siti Armiza Mohd Aris, Mahfuzah Mustafa, Nazre Abdul Rashid, Noor Hayatee Abdul Hamid, laboratory technicians and staffs of the faculty who are always working in group and willingness to assist each other during research work.

I would like to express my gratitude as well to my employer, Universiti Malaysia Pahang and Ministry of Higher Education for their sponsorship and financial support for my three and half years of research work in Universiti Teknologi MARA.

I am truly indebted to my proofreader, Associate Professor Shaikh Nasir Shaikh Ab Rahman, who had assisted me to improve the quality of the writing skill for this manuscript.

Finally, the thesis is dedicated to my father, my mother, my wife, my daughters and sons for their outstanding support and patience which giving me high motivation to complete this thesis.

TABLE OF CONTENTS

CHAPTER FOUR: METHODOLOGY

CHAPTER FIVE: RESULTS AND DISCUSSIONS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

Symbols

c	Number of cluster
e	Error in confusion matrix
J	Objective function of FCM
k	Classification index / Sequence counter
log	Logarithm
ln	Natural logarithm
n	Time index
p	Statistical significant value
r	Pearson correlation factor
T	Sampling period
v	Cluster center
x	Data point
Σ	Summation of sequence
δ	Delta band
θ	Theta band
α	Alpha band
β	Beta band
ω	frequency
μ	Degree of membership

LIST OF ABBREVIATIONS

Abbreviations	
ANN	Artificial Neural Networks
ANS	Autonomous Nervous System
ANOVA	Analysis of Variance
AR	Asymmetry Ratio
BIS	Bi-spectral Index
DFT	Discrete Fourier Transform
DWT	Discrete Wavelet Transform
EC	Eyes-Closed
EO	Eyes-Open
EEG	Electroencephalogram
EOG	Electrooculogram
ECG	Electrocardiogram
EMG	Electromyogram
ESD	Energy Spectral Density
FCM	Fuzzy C-Means
FKM	Fuzzy K-Means
FFT	Fast Fourier Transform
GSR	Galvanic Skin Response
HDRS	Hamilton Depression Rating Scale
IQ	Intelligent Quotient
k-NN	k-Nearest Neighbor
LOO	Leave-One-Out
MANOVA	Multivariate Analysis of Variance
MSE	Mean Squared Error
PSD	Power Spectral Density
PSS	Perceived Stress Scale
RER	Relative Energy Ratio
SC	Spectral Centroids
SE	Spectral Entropy or Shannon's Entropy

SRI	Stress Response Inventory
VFMSG	Voltage From Marconi Signal Generator
VRBG	Voltage Received By g-MOBIlab

CHAPTER ONE
INTRODUCTION

1.1 RESEARCH BACKGROUND

Stress is one of the major health issues where too much stress may lead to depression, fatigue and insomnia. Regardless of weight and type of stress, it affects human lifestyles and work performance. Stress can either be positive (eustress) or negative (distress) [1-3], and the prominent definition of stress is the failure of the human body to challenge stressors or stress factors mentally, physically and emotionally [3-5]. Among the major stressors are high workloads, noisy working area, improper sleep, unfinished work, fear of something and conflict in family. Stress disturbs the balance of sympathetic and parasympathetic level in the human Autonomous Nervous System (ANS), resulting in the release of stress hormone (*cortisol*) and leads human to experience negative stress (distress) such as depression, anxious, angry, fatigue and frustrated.

Researchers have introduced qualitative and quantitative methods to detect stress, the former method normally employs self-report questionnaires while the latter method analyses human physiological signals [6]. Physiological signals such as Electrocardiogram (ECG), Electroencephalogram (EEG), Electromyogram (EMG), Electrooculagram (EOG), Galvanic Skin Response (GSR), Skin Temperature, Blood Volume Pulse (BVP) and respiration rate can be utilised to identify stress. Meanwhile, Cohen's Perceived Stress Scale (PSS), Stress Response Inventory (SRI) and Hamilton Depression Rating Scale (HDRS) are widely used self-report questionnaires on stress. In spite of the various methods implemented to detect stress and provide a solution to overcome it, there is still lacking in methods for indicating the level of stress of healthy person.

Since stress is dependent on human emotions and linked with the Autonomous Nervous System (ANS), this research attempts to provide the stress index as a stress indicator based on cerebral activities and change of cognitive state in the human brain due to the putative stressors. EEG signals are selected to recognize human stress since the signals are generated from the electrical activity in the human brain due to the change in cognitive state or any stimulation to the brain. Furthermore, the use of EEG signals to measure human stress is more precise and reliable than the self-report questionnaires and other physiological signals [1, 7-8]. The analysis of the features obtained from the EEG signals due to stress can lead to the recognition of stress level.

An index is a numeral normally used as an indicator in the field of health, such as diseases, emotions, fatigue and stress, to trigger the necessary action. For instance, Bi-spectral Index (BIS) is used to index the level of anesthesia (unconsciousness) from 0-100 [9]. Meanwhile, the assessment and quantification of the level of fatigue has been determined from the analysis of EEG signals [10-11]. In addition, index is also used to recognize human emotion such as fear, sadness, peace and happiness [12]. The index can provide fast information to human for immediate action to be taken to overcome the health problem related either to diseases, drowsiness or emotions.

Even though many researches have been done to identify the stress level using physiological signals, researchers have yet to come out with a reliable index for stress level indicator using EEG signals from healthy people. Stress index can be used to indicate the seriousness of stress regardless of the stresses so that appropriate action can be taken to overcome the problem. The index of an individual stress level can be used as a guideline to provide suitable treatment and the good accuracy of the developed non-linear model can predict the tendency of an individual to have stress even though he or she is in good health.

In this research, using intelligent signal processing techniques, the characteristics of EEG signals and its features that might relate to stress are studied, identified and classified in order to construct a reliable stress index. Stress features in the EEG signals can be extracted using feature extraction techniques and the features can then be classified and clustered using k-Nearest Neighbor (k-NN) classifier, Fuzzy C-Means (FCM) and Fuzzy K-Means (FKM). The results from these classifications and clustering are used to assign stress index.

1.2 PROBLEM STATEMENT

Various methods and health indices have been implemented by researchers to recognize human stress level either using physiological signals or psychoanalysis tests, however a realiable stress index to indicate the stress level of a healthy person has yet to be produced. Whereas the stress experienced by an unhealthy person can be easily detected, this is not the case with a healthy person. In spite of the various stimulation techniques conducted on the human brain using human physiological signals, previous researchers have not come out with a reliable stress index to describe the stress level of a healthy person. Thus, a reliable biological indicator to indicate the stress level of healthy person is required.

1.3 RESEARCH OBJECTIVES

The main objective of the thesis is to develop a new technique to determine and classify human stress index using non-parametric analysis of spectral centroid on the ratio of the energy spectral density of the brain signals. To achieve this objective, the following sub-objectives were implemented.

- To construct signal processing technique that has the capability of identifying the stress features from the EEG datasets in terms of asymmetry ratio, energy ratio, asymmetry ratio and spectral centroid of the energy spectral density of EEG signals.
- To measure the classification and clustering performance for the group with stress and non-stress features using k-Nearest Neighbor (k-NN) classifier, Fuzzy C-Means and Fuzzy K-Means (FKM).
- To assign stress index based on the results of spectral analysis and classification of the selected EEG features.

1.4 RESEARCH SCOPE

To realise the objective of the research, all activities and experiments related to the research were conducted according to the following scope of research.

1.4.1 EEG Data Collection and Re-generation of the Corrupted EEG Data

EEG data were collected at the Research and Development Laboratory for Human Potential, Universiti Teknologi MARA (UiTM), Shah Alam, Selangor. The Laboratory is a controlled room that is free from environmental noise in order to minimize the interference of noise during EEG measurement.

The data were categorized into four groups: Group 1 consisted of 50 EEG data with Eyes-Closed (EC) state, 50 EEG data with Eyes-Open (EO) state (after answering IQ test question), 40 EEG data with EC state (before performing horizontal rotation), and finally, 40 EEG data with EC state (after performing horizontal rotation). Thus, a total of 180 data comprised from 140 subjects were employed in this research. However, 37 corrupted EEG data were replaced with new EEG data using EEG data re-generation technique in order to have equal representations of each experimental group. The EEG data were taken from the subjects with the age ranges from 20 to 50 years old. The selected subjects are required to be in healthy condition, non-smoking, not consuming any drugs and not undergoing any medication. The declaration of the health condition must be done prior to EEG measurement (refer to Appendix A). Since all the experiments that were conducted in this research involve human beings, written permission was put forward to every subjects to obtain their approval before carrying out the experiments.

1.4.2 EEG Equipment Set-up and Measurement Protocol

The EEG equipment (amplifier) and electrodes were setup according to the standard manual for EEG montage before taking any measurement. Besides, the EEG equipment was buy-off or validated in term of voltage and frequency. The Prefrontal area of a subject’s brain (forehead) was chosen as measurement area. Meanwhile, the measurement time was set according to the cognitive states, which was Eyes-Open (EO) and Eyes-Closed (EC) states with duration of 10 minutes and 3 minutes respectively.

In order to stimulate the subject’s brain, the subjects were instructed to answer 20 Intelligent Quotients (IQ) test questionnaires based on the modified Raven’s Standard Progressive Matrices (SPM) that involves logical thinking. Writing and oral communication were not included in the test.

1.4.3 Analysis of EEG Signals

The EEG signals were measured using EEG electrodes and amplifier. Then, the signals were transferred to computer using Bluetooth and were captured using SIMULINK in MATLAB. The analysis of the signals were done in off-line manner using MATLAB. The statistical analyses of EEG signals were carried out using SPSS, Excel and MATLAB.

1.5 THESIS OUTLINE

The thesis is organized as follows:

Chapter 1 briefly describes the need of the research, which includes the definition of stress, stress detection techniques, physiological signals and stress index based on EEG signals. Then, the main goal, specific objectives and the scopes of the research are outlined.

Chapter 2 presents the related literature review of the research. It elaborates in detail the definition of stress, the relationship between stress and EEG signals, the existing techniques and their limitation in indexing stress level, spectral analysis of EEG signals, features extraction methods and, finally the classification of the features using k-NN, FCM, FKM clustering and Z-score technique to develop the stress index.

Chapter 3 describes the theoretical, formulation and mathematical calculations involved in the research, which include the calculation of EEG spectrum, energy, the ratio of energy, spectral centroids, entropy, Z-score, k-NN classification, cross-validation of k-NN classification and classification performance in terms of accuracy, sensitivity and specificity. The chapter also includes the design of the band pass filter and artifact removal techniques.

Chapter 4 explains the methodology to achieve the main objective of the research. It covers data collection, data re-generation, equipment set-up, equipment validation, measurement protocol, analysis of EEG signals, and extraction of stress features, k-NN classification, FKM and FCM clustering, Z-score and statistical analyses.

Chapter 5 presents the results of Chapter 4 in the form of tables and figures. The results of the confusion matrix, the results of classification, clustering, statistical analyses and correlation study are discussed. The chapter also includes the result of artifact removal and equipment validation. Finally, the table of stress index is produced and discussed. The chapter also discusses the verification of the assigned stress index using Z-score technique and range of the EEG data from each group.

Chapter 6 concludes the research work and offers recommendations for future work

CHAPTER TWO
LITERATURE REVIEW

2.1 INTRODUCTION

The research begins with an extensive literature review on the existing stress detection techniques based on EEG signals and their limitations, which include non-parametric analysis of EEG signals, extracted features that are related to stress, and classification techniques to index the stress level. The chapter also describes the existing technique in producing the stress index. The chapter starts with literature reviews on the generation of emotions and EEG signals, and is followed with the relationship between emotions and EEG signals. Here, the focus of the study is negative emotion, which is stress. Since stress is associated with emotion, the literature review also includes emotion recognition techniques. Next, the chapter discusses the literatures on existing technique in analyzing and detecting human stress level, while scrutinizing its limitations in stress detection system. Since the study involves the non-parametric technique (spectral analysis) in extracting features related to stress from EEG signals, literatures related to this technique is thoroughly discussed. Meanwhile, the parametric technique (model-based) is also discussed as a comparison to non-parametric technique. Next, the chapter explains literatures related to the classification technique of selected stress features found in the EEG database. Since the study uses the k-NN classifier, FCM and FKM clusters to classify and cluster the extracted EEG features, this chapter only explains literatures related to these classifiers and clusters which includes cross-validation technique to validate the performance of the classifier. Finally, the chapter discusses the literatures on analyzing the selected EEG features in statistical manner.

2.1.1 Brain and EEG Signals

The human brain consisted of the Frontal Lobe, Parietal Lobe, Occipital Lobe and Temporal Lobe, as depicted in Figure 2.1, and is divided into the right hemisphere and left hemisphere.

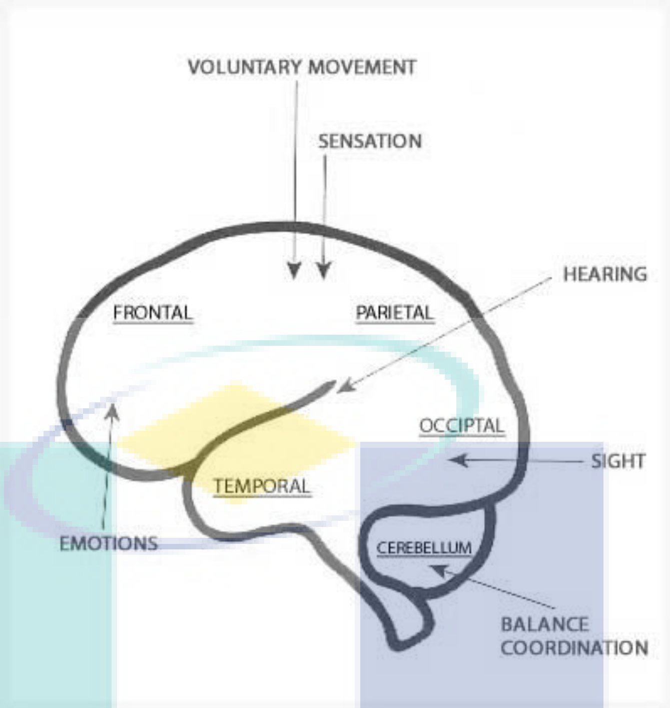


Figure 2.1: The Brain's Lobes and Their Functions [2-3,5]

The functions of the lobes are described in Table 2.1 [6-7,13-15]. The frontal lobe is the biggest lobe part of the human brain, where its main function is to generate emotions and cognition. Since this lobe is associated with emotions, this study focuses on this area of the human brain.

Table 2.1:
The Function of the Lobes in Human Brain [6-7,13-15]

Lobe	Function
Frontal	Emotions, memory, cognition
Parietal	Voluntary movement, sensation
Occipital	Sight, vision
Temporal	Hearing
Cerebellum	Balance coordination

Human emotions can be recognized by analyzing the electrical activity in the frontal lobe region of the brain, where the electrical activity will be generated by the change in cognitive state or any stimulation to the brain. This neural activity can be captured using Electroencephalogram or Electroencephalography (EEG) [7, 16]. The brain electrical signal recorded by EEG is called EEG signals or brainwaves. The EEG signals are non-linear, dynamic and random with a very small amplitude in electrical potentials ranging from 10 to 100 microvolts [7-8]. The characteristics of EEG signals are described by frequency and amplitude as shown in Table 2.2. The signals can be categorized into Beta band, Alpha band, Theta band and Delta band, the band represents the cognitive state of alert, relax, light sleep and deep sleep, respectively. In terms of frequency and amplitude, Beta band has highest frequency but lowest amplitude, while Delta band has lowest frequency but highest amplitude.

Table 2.2:
The Characteristics of EEG Signals[7-8,16]

EEG Sub-bands	Frequency (Hz)	Amplitude (Microvolt)	Cerebral Activities
Beta Band	13-30	Lowest	Alert or working state
Alpha Band	8-13	Low	Relax or eyes-closed
Theta Band	4-8	High	Drowsy or light sleep
Delta Band	0.5-4	Highest	Deep sleep

Researchers have found that emotions generated from the front side of the cerebral cortex can be categorized into two categories, which are positive emotions and negative emotions. Emotions can be modeled using Russell's two-dimensional Affective Space model, where the X-axis represents Valence (positive and negative emotions) and Y-axis represents Arousal (active and passive levels of emotions) as illustrated in Figure 2.2 [4,12].

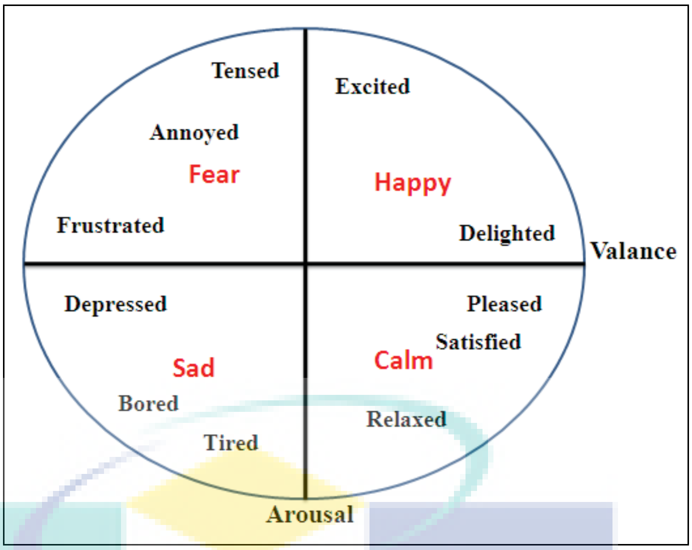


Figure 2.2: Russell's Affective Space Model [4,12]

Based on the model, there are four basic emotions that represent the positive and negative emotions. The emotions are Happy, Calm, Sad and Fear, which are scaled by valence and arousal [4, 12, 17], and stress falls into negative valence and active (positive) arousal. Stress might closes to *Tensed* under quandrant of Fear as shown in Figure 2.2. As discussed in Chapter 1 (Introduction), stress can be classified into positive stress (eustress) and negative stress (distress). Positive stress is vital in human daily live since it capables to improve or lead human to obtain good lifestyle. For example, human needs to has a positive stress or motivation to obtain a dream car, to take an important examination, to success in job interview, to purchase the first house, to deliver a baby, to get promoted at work, to graduate from college and others. Meanwhile, stress might also goes to *Depressed* under quandrant of Sad if human body unable to challenge stress and it will absolutely disturb human health [4, 12, 17].

EEG signals can be used to identify and classify types of emotions [18-20]. Researchers have introduced various techniques using EEG signals to recognize human emotions and the most popular technique is Frontal EEG Alpha Asymmetry [21-23]. In this technique, emotions are categorized based on the pattern of brain electrical activity (activity of Alpha band) at the right and left sides of the frontal lobe due to motivational approach versus withdrawal. Negative emotion, such as withdrawal or depression, is associated with right frontal activation or left frontal inactivation. Meanwhile, positive emotion, such motivational approach or happiness is associated with right frontal inactivation or left frontal activation [23-28]. The right and left activation is indicated by a difference score using Alpha asymmetry ratio [28-

30] and the method has been widely used in Psycho-physiology field to assess emotions.

Another technique that has been used to recognize emotions is the analysis of EEG signals measured at frontal and temporal lobe using Wavelet Entropy. The technique is able to classify two types of emotions (Calm and Excited) with 73.25% accuracy. Other than frontal asymmetry and entropy techniques, wavelet transforms and fractal dimension are widely used techniques to recognize emotion. In addition, features from EEG signals can be extracted using Discrete Wavelet Transform (DWT) technique with a classification accuracy of 83.26% [31]. The fractal dimension technique was introduced to recognize brain state in terms of emotion and the level of emotion [32]. In addition, the technique was used by the researchers to classify EEG signals that correlate with chronic mental stress [33].

Other devices that can be used to detect human emotions include Blood Volume Pulse (BVP), Electrocardiogram (ECG), Electromyogram (EMG), Electrooculogram (EOG), Galvanic Skin Response (GSR), respiration rate and skin temperature [34]. For instance, the physiological signals such as BVP, GSR and pupil diameter were used to detect negative emotion of computer users [34]. However, since the signals are generated from the frontal lobe of the human brain, EEG is most preferred methodology to detect and measure human stress (negative emotions). Researches have shown that the frontal lobe is a suitable location to measure stress based on the analysis of the EEG signal of that region. Besides, the increase in the ratio of Beta over Alpha power in frontal lobe can be used to indicate the strength of emotions [19]. An emotions recognizer has been developed to recognize stress at the frontal lobe with 83.2% accuracy [35].

Other than the brain lobes, the right and left hemispheres of the human brain also play vital role in studying human emotions. The cognitive functions of the right and left hemispheres are different, as described in Table 2.3 [36-38]. The cognitive functions of the right hemisphere are on creativity, intuition, perceiving and remembering, while that of the left hemisphere are on analytical, logical, language and speech. Previous researches have shown that the ability to use cognitive functions of both hemispheres could lead to a better lifestyle and health [37-40]. Thus, it is crucial to study cerebral activities on both hemispheres of human brain regardless of the type of emotions. The cerebral activity of the right hemisphere is associated with negative emotions, while that of the left hemisphere is associated with positive emotions [20-22, 40].

Table 2.3:
The Brain Hemispheres and their Cognitive Functions [36-38]

Right Hemisphere	Left Hemisphere
Creative	Analytical
Imaginative	Logical
Intuitive	Repetitive
Conceptual	Organized
Big Picture	Details
Heuristic	Scientific
Empathetic	Detached
Figurative	Literal
Irregular	Sequential

2.1.2 Stress and EEG Signals

Researchers have discovered that stress is associated with human emotions, either negative emotion (negative valence and active arousal) or negative psychological state. Nowadays, chronic stress disturbs human lifestyle and might lead to major health issues such as depression, drowsiness, mental disorder and sleep disorder. For instance, mental disorder caused by stress might lead to brain-related diseases such as Alzheimer, epilepsy, schizophrenia and dementia while depression might cause a person to commit suicide [41-42]. However, there is no single definition of stress among individuals as different individuals treat or react to stress differently. In the field of psycho-physiology, stress is defined as the disturbance in the balance of sympathetic and parasympathetic activities of the Autonomous Nervous System (ANS) due to the stress factors or stresses [43-46]. Stress occurs when human resist putative stressors emotionally, mentally and physically with the release of *Cortisol* in the human body [41, 47]. The putative stressors or processive stressors can be defined as the stress factors that most frequently affect human lifestyle such as high workload, noisy working environment, taking examinations, lost job, death of family members, getting exhausted and improper sleep [41, 48-49].

Researchers have divided the definition of stress into two categories, which are positive stress (“eustress”) and negative stress (“distress”) [1-3]. Positive stress is healthy and motivates people in their lives; for example, to obtain a dream car, to take an important examination, to success in job interview, to purchase the first house, to deliver a baby, to get promoted at work, to graduate from college and others. On the contrary, negative stress is harmful and most people suffered this of type of stress, which affect their lifestyle even though their health conditions are good. However, if their bodies are unable to resist stress for a long time, it will lead them to a chronic stress and finally, they might encounter dangerous diseases [1-3, 43-44, 49]. For instance, from analysis of EEG signals, researchers have found that subjects had suffered primary insomnia due to stress and depression [50]. Therefore, it is very crucial for humans to manage their stress.

Researchers have discovered that stress can be distinguished or categorized by different stress level such as high stress, moderate stress and low stress [51], while stress can be classified into physical, emotional and mental [5, 52-54]. There is a close relationship between the level of human stress and physiological signals. It has been observed that the relative Alpha power at Parietal and Occipital lobes of the human brain due to induced noise factor (physical and mental stress factors) was reduced, while that of Theta power was increased [49, 55].

2.1.3 Stimulation of Brain

As changes in cerebral activities cannot be detected without stimulating the brain, researchers have implemented several brain stimulation techniques. Such techniques include listening to music, watching pictures or movies or music videos, answering examination questions or stress questionnaires, performing IQ test or mental arithmetic tasks or mental stress test (*Stroop* test), and performing actual driving or driving simulation. If stress stimuli are applied to human, ANS will release the stress hormone *Cortisol* and cerebral activities will react to challenge the stress. Therefore, the usage of EEG technology after performing brain stimulation is vital in detecting features that might be considered as stress features.

There have been numerous researches on how to stimulate the brain to produce stress [56-58], such as the use of mental stress test on subjects to detect their stress level. However, the existence of stress can be evaluated by analyzing the features extracted from BVP, GSR and Pupil Diameter (PD) signals [34]. In addition, audio-visual and electromagnetic stimuli can be used to produce emotional stress in subjects, the effect of which is the lowering their cognitive activities [59-61]. Researchers have also discovered that audio-visual outperformed visual stimulus in stimulating frontal lobe of the brain in inducing emotional states [62-64].

Another popularly used visual stimulus to stimulate the brain is the International Affective Picture System (IAPS), in which subjects are shown various types of pictures or movies to generate positive (calm) and negative (excited) emotions [65-66]. The analysis of EEG signals after applying this stimulus is able to categorize emotions into positive and negative emotions [18]. Beside audio-visual stimuli, actual driving or simulation of driving are widely carried out by researchers to stimulate brain and generate mental stress and fatigue [67-68]. In this case, emotional stress can be assessed by recording and studying the frequency of eyes-blink during the driving session [11, 68].

Human affective states and their relation to emotions can be analysed by showing music videos to the subject, while recording the subject's EEG signals. The level of arousal and valence can be detected from the extracted features after spectral analysing the EEG signals [69-71]. Audio-visual stimuli are eminent in studying the relationship between brain and stress. The measured EEG and ECG signals produced in the frontal lobe of the brain of the stress group, as a result of applying audio-visual stimuli that acts as emotional stress tasks, can then be compared to that of the non-stress group [72]. Visual stimulation together some intervention (rotation bed) is used in this study but no audio stimulation is used. The stress stimulus used is the IQ test, a popular tool for measuring the intellectual level of the subjects. The brain of a subject is stimulated when the subject are requested to perform IQ test and undergo intervention while their brain signals were captured using the EEG device. The subjects are required to answer the IQ test questions, ranging from easy to difficult questions, in a stipulated time, and it is believed that the difficult questions will generate emotional and mental stress [refer to Appendix B]. Here, the difficult questions require subject to think and take longer time to answer the questions such as Question 14, Question 15, Question 18 and Question 20. The IQ test utilised in this study was modified from the standard IQ test questionnaires (Raven's Standard Progressive Matrices) [73]. As a reference, EEG signals were also recorded from a control group, subjects who were not undergoing the IQ test.

2.2 STRESS DETECTION TECHNIQUES

Stress is associated with negative emotions generated from the frontal lobe of the human brain, and researchers introduced numerous techniques to recognize and assess the level of stress. One widely used technique is based on physiological signals, while another is based on self-report questionnaires. The physiological signals, such as Heart Rate Variability (HRV), Skin Conductivity and EEG are used to detect stress levels of drivers [74]. The rate of eyes-blinking has been used to indicate the stress and drowsiness level during simulation of car driving [68]. In this study, beside eyes-blinking, HRV, ECG, EMG, skin conductance and respiration rate were recorded to determine the level of mental workload and stress of a driver while driving in a city and on a highway. The study was able to recognize low stress at the 100% classification accuracy, medium stress at 94.7% and 97.4% for high stress, while the slope of EEG linear regression was used to determine the human relaxation level [60, 63].

Other than the slope of EEG signals, its spectral and statistical analysis were extensively employed by researchers to capture the features that migh relate to stress. EEG spectral analysis of the Alpha band has been used to recognize human personality and characteristic [75-77]. In addition, EEG power spectra was used to determine acute stress, chronic stress and normal stress under a hot environment with classification accuracy of 96.67%, 97.17% and 98.50%, respectively [51].

Also, the asymmetry technique of EEG power at the frontal region of the human brain were introduced by researchers to discriminate the depressed subjects from healthy subjects [26-27, 42, 78-80]. Researchers have shown that, when detecting neural activity due to stress and fatigue, the change in EEG signals during stress and fatigue can be easily recognized using the entropy analysis of EEG signals instead of the spectral analysis technique [81].

Studies previously conducted have been able to segregate the stress level of a stress group from that of a non-stress group by studying their EEG signals, as the Beta band is much more active in the stress group compared to that of the of the non-stress group [49, 59]. A spectrum analysis using Short-Time Fourier Transform (STFT) can be applied to the EEG signals to determine the relationship between the human brainwaves and their mental tasks. In an experiment, one group of subjects was asked to be in a state of rest (relaxed and think of nothing), while another group was required to imagine the rotational movement of a cube after watching the image of the cube on the computer. It was found that the power of the Alpha band of the latter group has increased during the 10-second when the subjects were concentrating on the task and reduced when subject has lost attention [82].

The smart spaces system was designed to detect people who under stress and provide the necessary solution using *Moheet* architecture [83]. Too much stress or the failure human body system to challenge stress can lead one to fatigue or drowsiness, as in the case among soldiers who are exposed to physical and mental stress [84]. In a related study, the level of drowsiness among soldiers were determined by measuring their EEG, EOG and ECG, where the instruments were attached to their helmet and the physiological data were transferred to a computer using Bluetooth. Some researchers have induced or manipulated stress on their subjects by giving difficult mathematical tasks or difficult mental tasks as stressful stimuli, where their EEG Alpha and Beta power are monitored. Individual who has the lowest score of task performance might experience stress and need to provide proper Alpha biofeedback or neurofeedback in order to change from the state of stress and anxiety to that of rest and relax [44, 85-89]. Among human physiological signals, EEG signals are more reliable in measuring stress because it indicates the imbalance in ANS due to stressors [3, 41-43] as the signal potray the neural activities that result from the change of the cognitive state. Hence, EEG is an effective discriminator for chronically stressed individuals [90]. Assessment of stress can also be conducted using self-report questionnaires, such as Cohen's Perceived Stress Scale (PSS), Stress Response Inventory (SRI), Hamilton Depression Rating Scale (HDRS), Stress Self-Rating Scale (SSRS) and Profile Mood States (POMs) [91-92], where scores are used to determine the stress level before and after the subject undertook some tasks. PSS-10 questionnaires has been used to determine the degree of stressful in an individual and it has been found that the EEG signals and the results from PSS-10 is negatively correlated. PSS-10 is based on 10 questions and a 5-point frequency scale ranging from 0 to 40; a minimum score of 0 indicates the individual has never encountered any stress, while a total score of 40 indicates he has experience high stress level [93].

Meanwhile, 14-items Perceived Stress Scale (PSS-14) has been used to discriminate stressed and non-stress groups before EEG signals of the groups were measured and analyzed to verify the stress level of the groups [33]. SSRS has also been used to assess the stress level of subjects before undergoing stimulation tasks to produce an emotional state [59]. In addition, SSRS and SAM (Self Assessment Manikin) have been employed to assess the emotional state of subjects [3, 54, 94].

Based on the previous studies conducted on stress using physiological signals and stress questionnaires, it has been found that Skin Conductivity, Heart Rate Variability and EEG are suitable and reliable neurophysiological assessment tools to detect and assess stress level either from normal person, car or truck drivers and aircraft pilots [3, 11, 49, 74]. The change in magnitude of EEG rhythms (Alpha and Beta bands) due to the change of emotions and cognition at the frontal lobe can be utilized to indicate the existence of stress. However, the concentration of this study is on EEG signals and the related technologies, and the main focus is on the generation of stress index based merely on the analysis of EEG signals taken from the frontal side of the human brain.

2.3 METHOD OF ANALYZING EEG SIGNALS

Currently, EEG signals are being used in various fields such as meditation, diseases, reflexology, anesthesia and music. In the field of diseases, EEG signals were used and analyzed to detect epilepsy, sleep apnea, insomnia, Alzheimer, dementia and depression [95-97]. In addition, different types of music, such as soft or rock music, can stimulate brain and affect neural activities and is shown by the change in amplitude and frequency of the EEG signals.

Typically, the features from EEG signals can be extracted using two types of analysis, which are parametric analysis and non-parametric analysis [98-100]. Parametric analysis is a time-base analysis that requires minimum data recording, and it involves modelling base analysis, such as autoregressive (AR), moving average (MA) and autoregressive moving average (ARMA) models [100]. Researchers have evaluated EEG activity in young children with epilepsy by extracting the spectral features in the EEG signals that has been modelled using AR or ARMA [98]. However, the drawback of the parametric technique is that it is incapable of establishing the model property when different EEG signals are used, even though only a few EEG data is required. On the contrary, non-parametric method requires a lot of EEG data to obtain good spectral parameters [100-101]. Thus, there are advantages and disadvantages in using both parametric and non-parametric techniques, but the parametric analysis method is more suitable for stationary EEG signals, while non-parametric analysis is more competent and practicable in extracting features from non-stationary neural activities [100].

Another advantage of non-parametric method is that the spectral analysis parameters can be used to estimate the alertness level of human operators regardless of the length of the EEG signals being analysed [99, 102]. Non-parametric analysis involves spectral analysis (frequency domain and time-frequency domain) and wavelet analysis. Recently, spectral analysis of EEG signals via Fourier Transform, such as Short-Time Fourier Transform (STFT), has become a popular technique in applications involving EEG signals. This technique was famously known as time-frequency representation (TFR) due to its capabilities to convert one-dimensional signal (time domain) to two- dimensional signal (time and frequency domains). With STFT or TFR, also known as spectrogram or windowed Fourier transformation, short windows are applied at high frequencies while long windows are applied at low frequencies, and the EEG features could be localized in time and frequency [49, 103-105]. STFT is a very effective technique in decomposing vast non-stationary EEG signal. As EEG signals are weak and highly non-stationary random signals, the spectral components of the signals will be more effectively decomposed using the combination of the STFT with wavelet transformation [106]. The classification of motor imagery EEG has been achieved using wavelet transformation [107].

Time-frequency analysis using wavelet transformation has become a powerful tool in the extraction of features from EEG signals as it is able to identify and localize the transient features of the signals, and simultaneously, produces the wavelet coefficients that represent the distribution of energy in the signals. However, the drawback of the technique is that it is dependent on the selection technique of the most suitable and practical mother wavelet [20, 31, 106-109]. Kernel Density Estimation is another non-parametric technique used by researchers to extract EEG features by calculating density of the data points [110].

When the spectrum of EEG signals are analysed using Discrete Fourier Transform (DFT), the time-based signals will be converted to frequency-based signals, where power and energy in specific frequency bands are revealed based on the neural activity of the brain [99-100, 111-112]. The generated power obtained from the analysis is called Power Spectral Density (PSD) and the energy under the area of PSD curve is called Energy Spectral Density (ESD) [24, 61, 112]. The power spectrum of EEG signals represents the power at each frequency band of the signal, hence are known as Delta power, Theta power, Alpha power and Beta power, whose the magnitude of the power density is denoted in ($\mu V^2/Hz$) [24, 28]. EEG power spectrum has been used to estimate the stress level of subjects after performing Mental Arithmetic Task (MAT) and simulation of car driving [68, 71]. Moreover, the EEG power spectrum has been applied in recognizing depressed male patients with 91.3% classification accuracy [42]. In an acupuncture treatment, the power spectrum of EEG signals was applied to monitor Alpha and Theta, and it was found that the highest power was seen at EEG rhythms of Theta and Alpha indicating an increase in the relaxation level of human brain [113]. The technique was also used by a researcher to extract features from EEG signals captured from the subjects under mental stress [33]. In addition, the power spectral density of the EEG signals obtained from the subjects after performing Audio Vigilance Task (AVT) able to provide features to classifiers to classify mental fatigue of the subjects into several levels [114]. The negative and positive emotions can be clearly identified using the features obtained from the time-frequency analysis of EEG signals [101], as the energy contents buried in the signals is revealed from the analysis [115]. These literatures confirm that EEG power spectrum is very useful in identifying human characteristics or cognitive states.

Wavelet transformation and entropy can also be employed as a spectral analysis method for detecting stress. Discrete wavelet transform was employed to extract EEG features in order to recognize positive and negative emotions, such as happiness, disgust and fear [94], while wavelet packet energy was used to reveal features related to mental fatigue [85, 116]. The wavelet transformation technique has also been widely used in emotion recognition. However, this method is less preferable by researchers in analyzing EEG signals to recognize emotion, probably due to the difficulty in selecting suitable mother wavelet and decomposition level.

Another method of analysing EEG signals is to employ spectral entropy to detect irregularity patterns in the signals, where the spectral entropy is obtained after applying entropy to the power spectrum of EEG signals. There are three types of entropy, which are Renyi's entropy, Shannon's entropy and Tsallis's wavelet entropy and entropy is applied to detect the abnormalities of the distribution of energy in EEG signals. Entropy has been used to detect changes in cerebral activities due to stress and fatigue [117]. In studies conducted on subjects with mental stress and epilepsy, it was found that their EEG spectral entropy was basically lower than that of subjects without stress and epilepsy [95, 117-118]. Also, the entropy analysis of EEG signals was used in finding abnormalities of the EEG signals caused by human emotions [31, 65]. Spectral analysis has also been employed to quantify the spectral complexity during mental tasks (imagined right and left hand movement). In addition, the stress caused by mental workload can be identified by the statistical and spectral analysis of their work stress [119-120]. In this study, the ratio of the EEG power spectrum of all frequency bands are spectrally analysed in order to determine the stress pattern of the EEG datasets.

2.4 ARTIFACTS REMOVAL

The raw EEG signals are affected by artifacts that come from various sources, and in terms of physiological signals, EOG (eyes-movement or eyes-blinking), EMG (body movement) and ECG (Heart rate) are considered artifacts or biological noises. Among them, EOG is the major artifacts due to the characteristics of the signals which have frequencies ranging from 0.1 Hz to 38 Hz and amplitude greater than 100 microvolt [121-128].

There are many widely used methods to remove the noise and obtain the noise-free EEG signal, such as by applying Adaptive Noise Cancellers (ANC) but this technique requires reference signal [7, 121-122]. In addition, the EOG signals can be separated from the EEG signals using Independent Component Analysis (ICA) and Principal Component Analysis (PCA) [123-124]. Data adaptive de-trending method using a bivariate extension to an empirical mode decomposition (BEMD) [125] and wavelet decomposition using threshold wavelet coefficient have also been employed to segregate EOG signals from the captured EEG signals [126]. Noise removal through regression method in time and frequency domain [127] and blind source separation (BSS) are other popular techniques preferred by researchers; with BSS, its algorithm automatically detects the EOG signals and completely remove the signal without disturbing the content of original signal [112, 124].

Another technique discovered by researchers was the use of Recursive Least Square (RLS) adaptive filter with ICA to separate ocular artifact from EEG signals, in which case the EOG signals were measured and acts as a reference signal to be used in adaptive filter, and the signals were then removed using ICA technique [124]. However, the filtering techniques introduced by the researchers to remove artifacts are quite complex.

EEG signals can be contaminated by various noises originating from various sources such as from biological signals, instrument sensors and the environment. In this study, the main concern is the EOG signals originating from biological signals, where the EOG signals is simply removed from EEG signals by applying the threshold value of ± 100 µV to the amplitude of the signals, and any raw data with amplitude outside the threshold value will be rejected. Meanwhile, noise from instrument sensors was removed by setting the proper impedance to the sensors during EEG measurement. On top of that, the experiments were carried out in a room with controlled environment.

2.5 EEG FEATURES

The features from EEG signals can be extracted using either parametric or non-parametric methods. The features are mean amplitude, root-mean-square amplitude, standard deviation, wavelet sub-band entropy, spectral peak power, spectral peak frequency, spectral entropy, power ratio, relative power ratio, power asymmetry, spectral centroid, bandwidth, zero-crossing rate and spectral roll-off frequency [128-129]. Since this study only focuses on non-parametric methods of analyzing EEG signals to secure features related to stress, the chosen EEG features to establish stress index from the group of EEG datasets are relative power ratio, the asymmetry of power ratio, spectral centroids and spectral entropy. The selection of these features will be discussed in this section and section 2.6 in this chapter.

Spectral centroids, among the eminent features in analyzing EEG signals, is widely used in speech and recognition system to detect the dominant frequency in EEG signals [130-134]. Spectral centroids is more dominant than Mel Frequency Cepstral Coefficients (MFCC) technique in extracting audio and speech features as it is more resilience to noise and is able to identify the dominant frequency from noisy speeches [134]. Spectral centroids has been used as one of the EEG features to be an input to a classifier to discriminate healthy subject from epileptic subjects [103]. In applying spectral centroids, the audio or speech data need to be converted to power spectrum first and then filtered to the number of frequency bands. In this study, spectral centroids were applied to the relative energy ratio of EEG datasets and the features obtained were selected as an input for the classification process.

Another method that can be employed to detect features in EEG signals is known as the asymmetry technique. In previous studies, researchers have employed asymmetry technique to determine the dominant Alpha or Beta power from brain hemisphere [39-40, 96-97]. In this study, the asymmetry power ratio is selected as EEG features due to its effectiveness in detecting changes in neural activities in the left and right hemispheres after the brain has been stimulated or after performing mental tasks. However, this study involves a combination of several techniques, which are asymmetry, normalized asymmetry, energy or power ratio and spectral centroids. The power and energy of the Delta, Theta, Alpha and Beta bands on both hemispheres are calculated before applying asymmetry technique to determine the dominant power or energy either in the left or right hemisphere of the human brain. Another technique is the application of the entropy on EEG power spectrum. Spectral entropy can be used to verify the stress pattern from the spectral analysis of EEG signals. The ratio of Alpha power at frontal and parietal lobes of the brain are the preferable features to measure the load due the increase in mental tasks [135-136, 198-199].

2.6 ASSIGNMENT OF STRESS INDEX

It is very crucial to come out with a numerical indicator to represent the human stress level due to the increase cases of chronic stress which include healthy people. Therefore, various stress indicators using physiological signals have been designed and introduced by researchers [137-139]. One such example is the Psychological Stress Index (PSI) using normalization of the Heart Beat Interval (HBI) and the amplitude of Plethysmograph (PPG), which has been used to assess the human mental condition [137-139]. In addition, Heart Rate, Skin Conductance (SC) and EMG were used to determine three different stress levels of drivers [74]. EEG signals were also used to measure the different drowsiness level of drivers [10, 116, 140]. Meanwhile, Bispectral Index (BIS) with a scale from 0 to 100 were generated from EEG signals to determine the level of Anesthesia (unconsciousness) [9]. Also, physiological index was created from the difference between the baseline interval of EEG band power and test interval of EEG band power in order to measure cognitive load [136, 141]. Index created from human biological signals really work as an indicator and it very useful in trigerring necessary action. Therefore, a reliable physiological indicator is needed to indicate the degree of the stress regardless of the health status, whether good or bad. With the aid of physiological indicator, a physician, clinician or biomedical researcher can take proper action to those under stress.

An index can be defined as a numerical value utilized as an indicator to trigger the necessary action. It has been applied in various fields especially in the field of health such as emotion, brain-related diseases, depression, stress and fatigue. Different indices have been assigned to different type of emotions such as fear, sad, peace and happiness [12, 94] but studies in consuming index for detecting stress are still few. The available stress index based on EEG features are summarized in Table 2.4, which elucidates that all researchers had applied the non-parametric method to EEG signals to produce indices for the any change in the brain activity due to the stimulation of the brain or performing any tasks mentally or physically. Since the objective of this study is to generate stress index for healthy person using EEG signals, and not from self-report questionnaires, this chapter will only discuss the non-parametric techniques that are used in producing stress index using the extracted features from EEG signals.

TABLE 2.4:
Summary of Existing Research on Stress using EEG signals

Ref. No.	Authors	Methods	Results
1	S. Handri *et al.*	EEG, ECG & Skin Temperature & k-NN Classification	Two levels index for Mental Stress (Low & High)
51	R. K. Sinha	EEG Power Spectra & Neural Network	Three indices: Acute, Chronic & Normal Stress; based on heat
60	M. Teplan	Dynamic Bayesian Network & EEG	Model to estimate human stress level
78	H. Hinrikus *et al.*	Relative difference in EEG Power Spectrum at two frequency bands	Index for depression – SASI (EEG Spectral Asymmetry Index)
117	Y. Tran *et al.*	EEG Power Spectrum and Entropy	Indicator for fatigue and stress; based on Entropy
135	A. Holm *et al.*	EEG power of Theta & Alpha ratio	Index to estimate cognitive workload, mental fatigue and stress

139	A. Nassef *et al.*	Ratio of Theta power at Frontal area with Alpha power in Parietal area	Task Load Index (TLI)
142	I. Jraidi *et al.*	Multi Agent System on Learner's brainwaves	Predict stress level variation of Learner
143	T. Yamakoshi *et al.*	Physiological signals (EEG,ECG,EOG, Respiration & Blood pressure)	Index for Driver Activation State (DAS) for monotonous driving
144	C. Berka *et al.*	EEG signals	EEG Index for Alertness, Cognition and Memory for mental tasks

There are various physiological indices produced by researchers, either using biological signals or non-biological signals. However, this study will tackle the stress index using a combination of relative energy ratio, entropy and spectral centroids in order to produce a reliable stress index. In the process of generating stress index, this study also evaluate asymmetry technique as a comparison to relative energy ratio in order to search for the best technique to produce reliable stress index.

2.7 SELECTION OF CLASSIFIERS

After obtaining the EEG features from non-parametric analysis of EEG signals, the features are then required to be classified in order to get the desired output. For instance, the classification of EEG transient events and features using rule mining and fuzzy has been used to detect epilepsy signals from the EEG signals [145]. Researchers have introduced various classifiers for classifying EEG features, such as Artificial Neural Network (ANN), Bayesian network, Decision tree, Fuzzy, Gaussian Mixture Model (GMM), k-Nearest Neighbor (k-NN), Linear Discriminant Analysis (LDA), Random Forests (RF) and Vector Support Machine (SVM).

A classifier is selected depending on the complexity and performance of the classifier in solving specific problems or classifying features. In a study, Fuzzy, k-NN and Bayesian network have been evaluated in diagnosing diseases, where it was found that Fuzzy outperformed k-NN and Bayesian in terms of high accuracy and sensitivity but selections of Fuzzy parameters were difficult [146]. Meanwhile, the combination of k-NN and LDA classifiers were employed to classify features from EEG signals in detecting human emotions with a prominent accuracy of 83.26% and 75.21% accuracy [31]. The performance of k-NN is dependent on the selection of the EEG features as the input to the classifier. The use of features from time-domain (cross-correlation) as input to k-NN have resulted in a classification rate of 97%, higher than that obtained when the features from frequency-domain (Power Spectral Density) are used [103]. In addition, it has been found that the performance k-NN in classifying human emotions is dependent on the number of centroids used [147].

ANN is a supervised learning algorithm that uses multi-layered perceptron (MLP) to minimize the difference between output and target values [7, 112]. In a study, ANN was used to discriminate the rate of mental task of subjects who were asked to perform five mental tasks, and the classifier has produced a classification rate of 80-86% accuracy [148-149]. ANN has also been used in classifying EEG features to detect human emotions and personal characteristics [33, 75-76, 150]. ANN has produced very high accuracy when applied in classification of stress in experiments conducted on rats that were exposed to heat, where the stress was classified into acute, chronic and normal [51]. In addition, ANN has been employed in the classification of EEG features from human emotions and achieving an accuracy of 96.43% in the predicted level of stress [4], and in the determination of the level of human vigilance using EEG features [151]. However, the drawback of using ANN classifier is that its parameters must be properly set so as to produce lower errors, where the selection of the parameters is determined by Casmean square error for every run. The process of analyzing EEG signals is time consuming even though the performance of the classifier is good. The most critical part in neural architecture is the selection of the number of hidden layer and numbers of neurons [108, 152]. The classification of three mental tasks (imagination of moving left-hand, imagination of moving right-hand and generation words of random letter) has been successfully implemented using three hidden layers feed forward ANN with a classification accuracy of 68.35 % [153]. In brain-related disease such as epilepsy, using ANN to classify EEG features extracted from the autoregressive modeling of EEG signals has successfully discriminated patients with epilepsy and patients without epilepsy with 92.3% accuracy [154-155].

In comparing ANN and SVM in terms of parameters setting, SVM requires more complex parameters. SVM has been widely used in classifying features related to diseases such as epilepsy and depression [7, 103]. In another research, SVM was employed to determine the subjects with epilepsy using the reduced EEG feature dimension obtained from the application of PCA, ICA and LDA on the features [98]. In SVM, discriminant plane or hyper-plane is created to separate datasets from different classes; the data points that were required to determine hyper-plane are called support vectors (SV). SVM uses non-linear mapping to present higher feature space dimension and search for the maximal margin between two classes from the hyper-plane [7, 33, 103, 156, 157]. In stress detection technique, SVM has been used to classify features from physiological signals, such as GSR, BVP, ST and PD, to determine the affective states of computer users, whether they are stressed or not, and the results of the experiment achieved 90.1% classification accuracy [34]. The combination of k-NN and SVM has also been used to classify subjects with chronic mental stress using features obtained from STFT, Fractal Dimension and Gaussian Mixtures of EEG Spectrogram (GMM) and the techniques were able to produce a classification rate higher than 90% [33]. On the contrary, LDA utilizes hyperplane to classify data into different classes with less complex mathematical calculation but this technique is difficult to employ when dealing with the nonlinear EEG data [158]. RF is nonlinear classifier and very effective in selecting huge features especially from EEG signals [114]. However, the classifier requires feature ranking criteria and it needs to be combined with Initial Feature Ranking scheme (INIT) or Recursive Feature Elimination scheme (RFE).

In this study, in order to support the supervised classification technique, the unsupervised classification techniques such as Fuzzy C-Means (FCM) and Fuzzy k-Means (FKM) are chosen to cluster the selected EEG features. Both FCM and FKM use the number of clusters to cluster the datasets contain feature-vectors. The performance of FCM is measured based on the membership degree between the cluster centers with the datasets, while that of FKM is evaluated on the distance of centroids in each cluster. FCM and FKM are widely used in clusters features and to obtain pattern from the datasets containing feature-vectors. For example, FCM has been used to cluster three types of human emotions (fear, happy, disgusts) after the EEG features were obtained using Wavelet Transform. Even though the number of clusters became an issue, the results of the study showed that FCM was prominent in the clustering of EEG features [159]. Moreover, both FCM and FKM were used to correctly classify six types of human emotions (anger, disgust, fear, happy, sad and surprise) [160].

Beside emotions, FCM has been used to classify EEG signals from EEG features extracted using ICA and Wavelet transform to recognize brain diseases, such as Epilepsy and Cerebral Palsy, with 95% classification accuracy [161]. In a study, Fuzzy was used to discriminate subjects with mental disorder from healthy subjects with 80% accuracy, with EEG features acted as an input to the Fuzzy [162-163], while in another study, FKM produced a good clustering performance to cluster the EEG features in detecting migraine [164]. FCM was also employed in detecting the level of stress in drivers using features from physiological signals such as ECG, EMG, SC and respiration rate. In FCM analysis, the stress value was calculated from the membership degree of EEG features and clusters, and it was found that the larger the size of the membership degree, the larger is the stress value [165]. Fuzzy expert systems are now becoming a good choice in detecting human stress, even though some parameters in fuzzy such as the selection of number of clusters remain an issue, as the selection of the parameters can be made based on the size of the datasets or data points. For example, the use of Fuzzy expert systems were able to detect stress at 99.5 % accuracy using features from ECG and GSR [166]. Besides detecting human emotions and diseases, FCM also can be used to track human activities. For example, FCM was used to classify the EEG pattern of human activity (before and after a meal, before and after smoking), where the EEG features were extracted using singular value decomposition (SVD) [167].

There are advantages and disadvantages in using classifiers to classify EEG features. Some factors that should be considered when selecting classifiers are the number of features, classification time, non-uniform weighting of the features, nonlinear map between inputs and output, unknown data distribution and monotonic convergence of the data. Beside the ratio of testing and learning vectors, various factors affect the classifier’s performance, one of which is the features themselves. In this study, k-NN, FCM and FKM are selected to classify and cluster the EEG features that might relate to stress. k-NN is selected due to its robust supervised learning algorithm and ease of use even though the k-value need to be varied depending on the number of features. Meanwhile, FCM and FKM have good performance in clustering human emotions.

2.7.1 Cross-Validation of Classifier

The EEG classification performance needs to be validated using cross-validation technique in order to reduce the bias created from the selection of training and testing ratio for EEG datasets [168]. From the various techniques currently available, in this study k-fold cross-validation and leave-one-out cross-validation are used to validate the k-NN performance. For k-fold cross-validation, k indicates the number of folds to validate the classifier, where the size of the fold is determined by the size of the datasets [168]. In 5-fold cross-validation, four folds are assigned for training and one fold is assigned for testing. It is impractical to use a big number of folds when the size of the datasets is small, and the popularly used number is 10-fold [169]. The advantage of using k-fold cross-validation is that all data will be employed for testing and training. In this study, the 10-fold cross-validation is selected as the size of the datasets is 180x4. Besides, the researches preferred to use 10-fold cross validation in validating k-NN performance [170]. Meanwhile, for the leave-one-out technique, one data will be selected for testing while the remainder will be selected for training.

2.8 STATISTICAL ANALYSES

SPSS (Statistical Package of Social Science) is a statistical analysis software package widely used by researchers to identify the significant differences of the selected features after securing the features from non-parametric analysis. For example, analysis of variance (ANOVA) and MANOVA (Multivariate analysis of variance) were used to reveal the significant difference in the selected EEG features obtained during monotonous driving [171-172]. ANOVA, a statistical method under mean analysis of SPSS [173], is an extension of t-test that can be utilized with a maximum of two independent datasets. It computes the significance values base on the mean and variance of datasets. ANOVA has been used to check the significant difference of the selected features obtained from EEG frontal asymmetry technique during cognitive task (mental arithmetic task) [29]. Also, ANOVA has been employed to obtain the significant difference of the mean scores of relative power spectra density ratio of various emotional states [59], to compute the significant differences between features obtained from EEG and ECG with *Cortisol* data [3], and to identify the significance of time domain features of EMG signals to classify four types of emotions; neutral, positive, negative and mixed [174]. Meanwhile, t-test method was used to examine the significance of the extracted features to recognize chronic mental stress [33].

MANOVA is the extension of analysis of variance that can effectively compute significant difference of more than one dependent variable. It requires dependent variables and covariates that are not correlated to each other in order to capture the unique variance, it has been found that MANOVA can compute and distinctly separate the features into two groups [175]. However, this study only applies ANOVA which is enough to discover the significant difference for the selected EEG features. Pearson correlation (scatterplot) is another statistical method that employed to search for linear relationship between two random variables. It can be used to indicate whether the correlation between the variables is strong or weak by calculating the correlation coefficient, r and the significant value, p. Those technique was employed to detect noise that produced by human head or body movement in measuring EEG signals [176]. In addition, Pearson's r was used to determine the non-linear correlation between EEG channels with its cognitive states (eyes-closed and eyes-open) [176].

In this research, in order to determine the significant value and correlation coefficient of the selected EEG features such as asymmetry and relative energy ratio, statistical analyses were performed using ANOVA and Pearson correlation. Before performing statistical analysis, the EEG datasets were checked for normality. The abnormality of the EEG datasets will affect the significant value and correlation coefficient of the EEG datasets. All the analyses were done using SPSS version 17. The statistical analyses are essential in confirming the significant value of the selected EEG features before the features can be used for the classification process. This process will assist classifier to produce good classification results, and the results of the analyses will be discussed in Chapter 5.

2.9 CHAPTER SUMMARY

Researchers have introduced various stress detection system using physiological or non-physiological signals. In addition, researchers have proved that physiological signals are much better in detecting stress, especially using ECG, EEG and Skin conductance. Meanwhile, the non-physiological method such as self-report questionnaires are imprecise. Since stress is part of negative emotions and originated from the frontal lobe of the human brain, EEG is more preferable and precise tool for capturing human stress. Therefore, using EEG signals, this research will tackle negative stress by producing a stress index that can be used to indicate the level of negative stress. Regardless of the type of stress, the natural response of a human brain will be registered by a change in power of brainwaves frequency components. In order to create a reliable EEG-based stress index, the analysis technique of EEG signals is emphasized.

Two techniques can be used to analyze EEG signals, which are parametric and non-parametric techniques. There are advantages and disadvantages of using parametric and non-parametric analysis in analyzing EEG signals to obtain features that might relate to stress, and the selection of the features plays a vital role in producing a reliable stress index. In this study, non-parametric analysis, namely, spectral analysis was selected, but the focus is on using Fourier Transform to convert the EEG signals from time-domain to frequency-domain (power spectrum). Furthermore, various features can be extracted from the EEG signals, either in time domain, frequency domain or time-frequency domain, however, in this study, the selected features are asymmetry ratio, relative energy ratio, spectral centroids and spectral entropy. These features have been found to be useful in determining different neural or cognitive activities in the brain. Hence, in order to implement a precise stress detection to produce a reliable stress index, this study propose to combine asymmetry and relative energy ratio with spectral centroids and spectral entropy. Based on previous researches on stress detection using EEG signals, the selection of the correct and reliable EEG features are most important in order to obtain a good classification accuracy.

For classifying the EEG features, k-NN is chosen due to the robust supervised learning algorithm and unsophisticated pattern classifer for spectral analysis of EEG signals, and FCM and FKM were selected to assist the classification. The k-NN was tested for robustness by validating the classifier performance using k-fold and leave-one cross-validation techniques. In order to figure out the significance and correlation of the EEG features, the statistical analyses are implemented using ANOVA and Pearson correlation.

CHAPTER THREE
FUNDAMENTAL CONCEPT AND THEORY

3.1 INTRODUCTION

This chapter details the theory and mathematical expressions that are used in this study to extract features using non-parametric analysis of EEG signals. As discussed in Chapter 2, non-parametric analysis can provide better analysis in term of feature extraction than parametric analysis when involving a lot of non-stationary cerebral activities. Since this study uses huge EEG datasets, non-parametric analysis is more preferable than parametric analysis. The analysis of the EEG signals involves the following steps: offset setting of the electrodes to capture the EEG signals, removal of EOG signals from EEG signals, filtering and categorizing of the EEG rhythms into Delta (δ), Theta (θ), Alpha (α) and Beta (β) bands, and the conversion of EEG signals from time-domain to frequency-domain using Fourier Transform technique. This chapter also discusses the theory related to Power Spectral Density (PSD) and Energy Spectral Density (ESD). In order to produce reliable features, the PSD and ESD values of the datasets are tested for normality. Since some datasets are corrupted, the corrupted data need to be replaced using data re-generation technique which will be discussed in this chapter.

Selection of EEG features, which are asymmetry ratio (AR), relative energy ratio (RER), spectral centroids (SC) and spectral entropy (SE), is carried out after the values of ESD has been obtained from the spectral analysis of EEG signals. Once the EEG features are selected and confirmed to be used to determine the stress index, then, the spectral entropy and Z-score analysis technique are employed to verify the stress index from the selected EEG features. Next, this chapter discusses the theory and mathematical expressions of the classifier and clustering techniques employed in this study, where the selected features become an input to the classifier. The theory and equation to validate the performance of the classifier is also discussed. Finally, this chapter will explain the theory of ANOVA and Pearson correlation since these statistical methods are employed to verify the significance of the selected features and the correlation between the cognitive states.

3.2 EEG SIGNALS PRE-PROCESSING

Due to the weakness of the EEG signals and the high tendency of the signals to be interfered by various noises or artifacts, the noises detection technique is utmost required. In this study, two channels are attached to the human forehead to capture the signals, and the first source of noise comes from the setting of the electrodes or sensors (instrument), while the second source originated from human neurophysiological signals. Since the measurement areas are close to the eyes, eyes-movement or eyes-blinking will be the major artifact and influences the characteristics of the EEG signals. Hence, the two types of noises is discussed in the following section.

3.2.1 Electrodes Offset

Several experiments were carried out to check the EEG readings captured by the electrodes, which were labeled as right and left electrodes, where it was found that there is a slight offset between the readings detected in Channel 1 (right electrode, Fp1) and Channel 2 (left electrode, Fp2). The procedure to validate the instrument and electrodes will be discussed in Chapter 4, and the results of validating the instruments are discussed in Chapter 5. This value is called "electrode offset" and is input into the MATLAB program before removing artifacts from the EEG signals.

3.2.2 EOG Removal

As stated in Chapter 2, the eyes-movement or Electrooculogram (EOG) signals are characterised with an amplitude greater than 100μV and frequency below 40 Hz [121-128]. Researchers have discovered that EOG signals have about similar characteristic with EEG signals in term of the range of the signal frequency. Consequently, the recorded EEG signals might consist of EOG signals, regardless whether the subject is in deep sleep, light sleep, relaxed or tensed. Other physiological signals, such as ECG and EMG signals, have less influence in affecting EEG measurement since the frequency of ECG signal is less than 1 Hz and that of EMG signal is greater than 50 Hz. Meanwhile, the non-physiological signal such as power line noise has a frequency above 50 Hz, where in the case an EEG device is unable to remove the noise, a notch filter will discard the noise [7]; a notch filter is not used in this study. Since the frequency range of EEG signals is from 0.5 Hz to 40 Hz, it will be affected only by EOG signals. Even though frequency becomes a key factor in recognizing artifacts, in this study, the amplitude of the artifacts are removed, as the corresponding frequency will be removed once its amplitude is removed. In actual implementation, the threshold value is set in the analysis program of MATLAB, such that any EEG signal with amplitude above +100μV and below -100μV will be rejected. Based on the literature review on the EOG signals, researchers have confirmed that the signals that were captured from the neural activities with amplitude above +100μV and below -100μV, is actually EOG signals, not EEG signals.

3.2.3 Band-pass Filter Settings

In this study, Finite Impulse Response (FIR) band-pass filter is selected to segregate the EEG signal components into its sub-bands. FIR filter is selected over Infinite Impulse Response [IIR] filter as FIR filter is more stable, has a linear phase response and is easier to implement compared to IIR filter [61, 112, 177-178]. Also, a window-based FIR filter is used in this study, and among window functions that can be used are Rectangular, Blackman, Hamming, Hanning (Hann), Gaussian, Bartlett and Kaiser. In designing of filters, filter order, type of filter, cut-off frequencies and window functions play a vital role to produce a required filter coefficient. The filter coefficient will then be convolved with the raw EEG data to produce filtered data. The filter order and its window functions provide smoother filtered data when the raw EEG data undergoes convolution through the filter. The filter order is calculated based on the frequency response of the specified cut-off frequency and sampling frequency. The major characteristic in designing the FIR filter is all frequencies of the filter is expressed in terms of normalized frequency and meet the cut-off frequencies of the filter specifications as stated in Table 3.1. Meanwhile, Figure 3.1 illustrates the plot of the filter's frequency response based on the filter specifications. This filter response must meet the filter specifications as stated in Table 3.1 in term of normalized cut-off frequencies, filter order, filter type and stop-band attenuation. In term of filter type, the plot of the filter's frequency response has confirmed the specified filter type in Table 3.1 which is band-pass filter. Next, this FIR band-pass filter must meet the normalized cut-off frequencies for the overall band, which are around 0.0039 to

0.2344 as specified by item no. 5 in Table 3.1. Also, the highest side-lobe (stop-band attenuation) of the filter should not exceed -53 dB, and as shown in Figure 3.1, the highest side-lobe is below -50 dB, which implies that the filter's specification is met [177-178]. The selection of the side-lobe or stop-band attenuation is based on the selected window as specified by item no. 8 in Table 3.1. Meanwhile, the filter order as specified by item no.7 in Table 3.1, is calculated from the window transistion width. If the frequency response of the filter fails to meet the specifications, the filter order need to be varied and the window function must be changed. In this study, a band-pass filter need to be designed in order to extract the frequency bands based on the characteristics of EEG signals as discussed in Chapter 2. Here, the band-pass filter is more practical and effective to filter the raw EEG signals with the range of the frequency from 0.5 Hz to 30 Hz.

Table 3.1:
Filter Specifications

No.	Filter parameters	Settings
1	Sampling frequency	256 Hz
2	Maximum signal frequency	256/2 = 128 Hz
3	Frequency bands	Delta = [0.5 - 4] Hz Theta = [4 - 8] Hz Alpha = [8 - 13] Hz Beta = [13 - 30] Hz
4	Normalized cut-off frequencies of frequency bands	Delta = [0.0039 - 0.0313] Theta = [0.0313 - 0.0625] Alpha = [0.0625 - 0.1016] Beta = [0.1016 - 0.2344]
5	Normalized cut-off frequencies for overall bands	[0.0039 - 0.2344]
6	Filter type	FIR, band-pass
7	Filter order	74
8	Type of window	*Hamming*

Table 3.1 elucidates the settings used to design the FIR band-pass filter and obtain the frequency response of the filter shown in Figure 3.1. The sampling frequency is set to 256 Hz in accordance with the sampling frequency of the wireless EEG amplifier (g.MOBIlab) and the maximum frequency 128 Hz, half that of the sampling frequency. Since the EEG frequency bands lies in the range between 0.5 Hz

and 30 Hz, the cut-off frequencies of the EEG frequency bands are computed by dividing the EEG frequency bands with half of the sampling frequency (maximum frequency) to produce a normalized cut-off frequencies of 0.0039 and 0.2344. The determination of the sampling frequency must follow the Nyquist rate theorem where the theorem has defined that the maximum input or analog signal frequency must be half of the sampling frequency. Meanwhile, the filter order of 74 as defined in Table 3.1 is selected because the normalized cut-off frequencies are met as shown by the graph of frequency response in Figure 3.1. In designing the window filter, *Hamming* window is selected in order to produce a good filter coefficient based on the response of the filter [177-178]. The specifications of the window is discussed in the next section of this chapter. The process to find the filter order and frequency response can be referred to Appendix E.

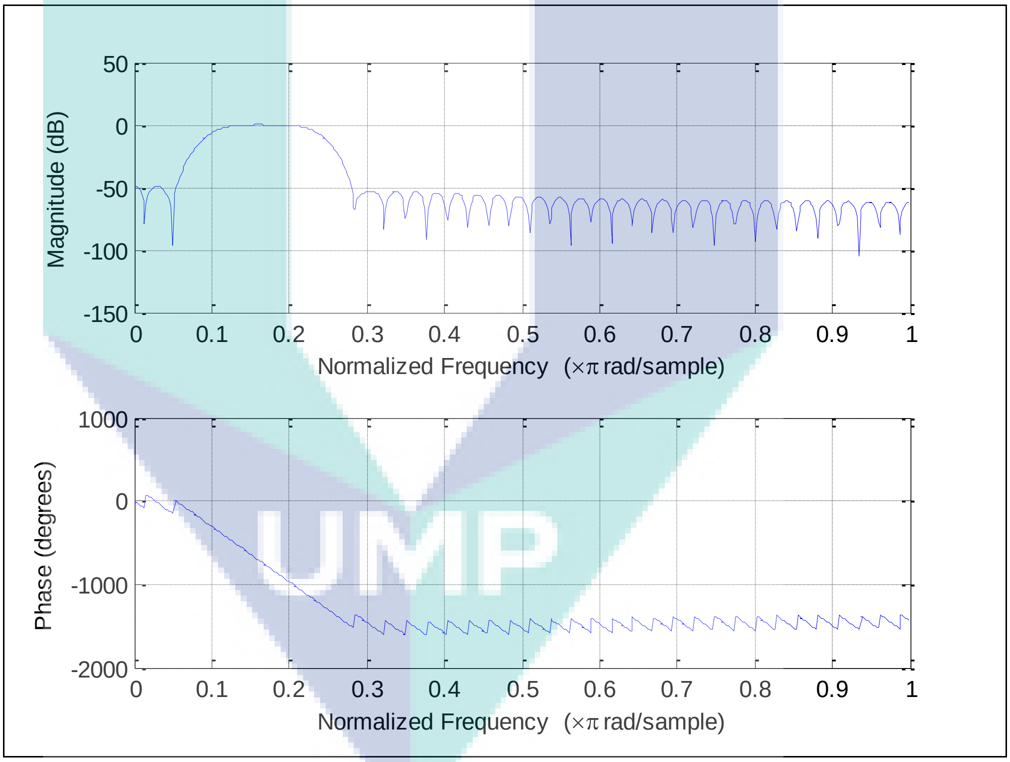


Figure 3.1: Frequency Response of the Design Filter

3.2.4 Window Specifications

The windows that can be used in filter designs are shown in Figure 3.2. From the figure, it is noted that the width of the *Hamming* window is wider than that of *Hanning* and *Blackman* windows. However, the *Hamming*, *Hanning* and *Blackman* windows have lower side lobes than that of the *Rectangular* window. Therefore, the

Hamming window is selected in this study based on the width of the window and the magnitude of the side lobes of the frequency response of the window. Due to the characteristics of the *Hamming* window, the use of this window in designing filter should result in good filter coefficient [16, 145-146]. The filter coefficient is used to capture the required EEG signals from the raw EEG signals during convolution process.

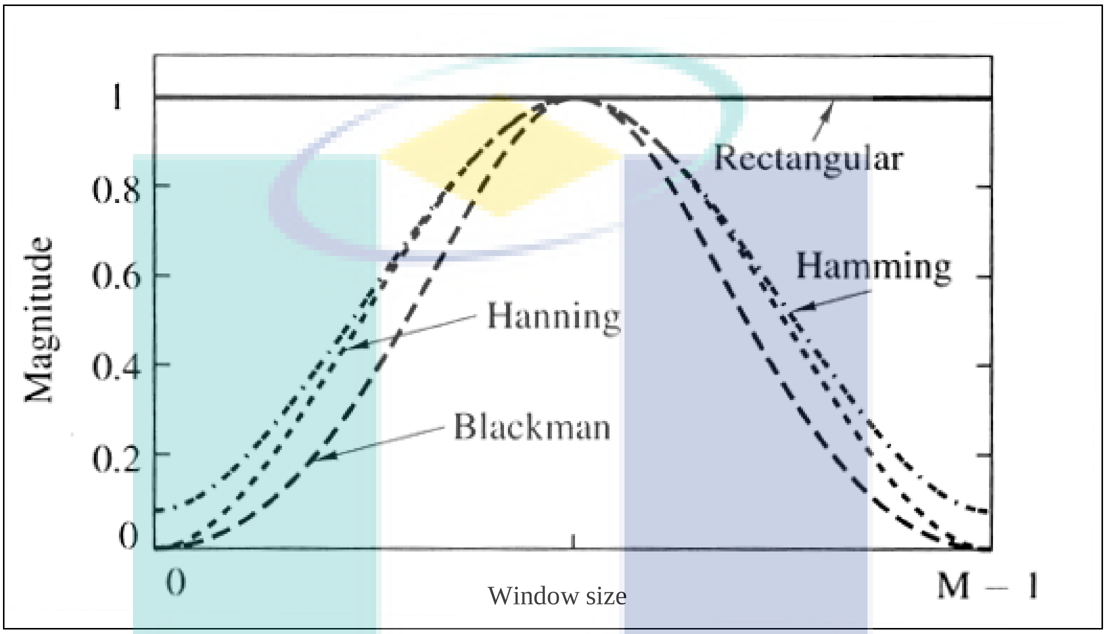


Figure 3.2: Type of Window Functions for Designing Band-pass Filter [146]

The specifications of the windows shown in Figure 3.2 are listed in Table 3.2, where N or M represents the filter length or order. Based on the selected filter order of 74 and *Hamming* window, the transition width will be 0.0446 by dividing the 3.3 with 74 (3.3 / N) and matched with the stop-band attenuation at -53 dB.

Table 3.2:
Window Specifications for Designing Band-pass Filter

Type of Window	Transition Width	Stop-band Attenuation
Rectangular	0.9/N	-21 dB
Hanning	3.1/N	-41 dB
Hamming	3.3/N	-53 dB
Blackman	5.5/N	-75 dB
Kaiser	2.93/N, 4.32/N, 5.71/N	-50, -70 and -90 dB

Table 3.3 describes the mathematical equations that were used to produce the windows, where the length of the windows is determined by the filter order. The formula also describe the size of the window which denoted by the filter order, N.

Table 3.3:
Mathematical Equations of the Windows

Type of Window	Mathematical Equations
Rectangular	$w(n) = 1, \ 0 \leq n \leq N$ 0, elsewhere
Hanning	$w(n) = 0.5 - 0.5\cos(2\pi n/N), \ 0 \leq n \leq N$ 0, elsewhere
Hamming	$w(n) = 0.54 - 0.46\cos(2\pi n/N), \ 0 \leq n \leq N$ 0, elsewhere
Blackman	$w(n) = 0.42 - 0.5\cos(2\pi n/N) + 0.08\cos(4\pi n/N), \ 0 \leq n \leq N$ 0, elsewhere

Figure 3.3 depicts the window obtained from the filter specification, and the size of the window is influenced by the filter order or length (N or M), in this case the length of the window will be 73. This means the raw EEG data will be filtered by a FIR band-pass filter with window functions having a width of 73. With the specified parameters, the filter produces characteristics that are similar to that of the *Hamming* window, as described by Table 3.2 and illustrated in Figure 3.2.

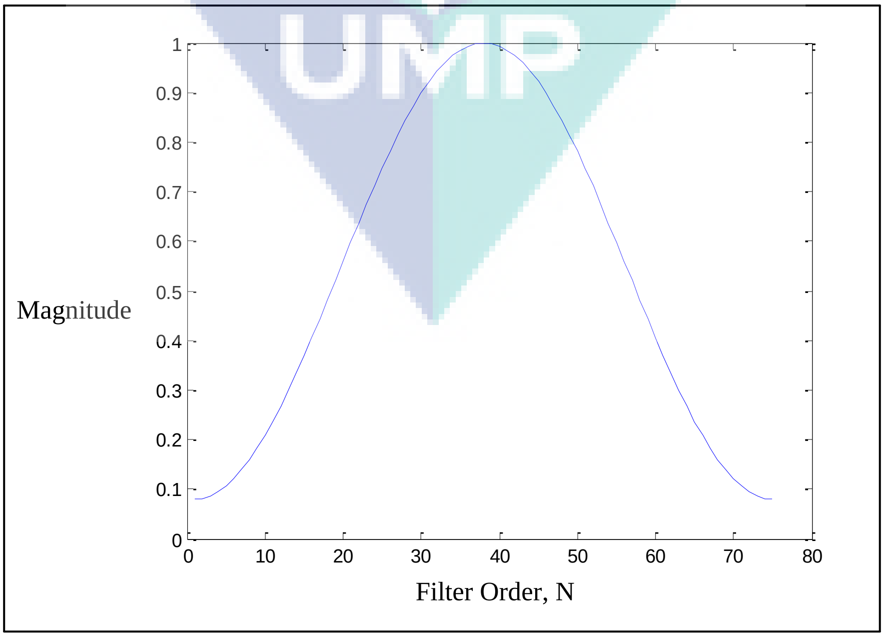


Figure 3.3: Hamming Window Produced from the Selected Window Specifications

3.2.5 Filter Coefficient

The frequency response of the N^{th} order FIR filter can be mathematically expressed by equation 3.1. In the equation, *h(n)* is defined as the filter coefficient, which will be convolved with raw EEG data to produce the filtered EEG data. $H(e^{j\omega})$ or $H(\omega)$ is the frequency response of the filter at the specified cut-off frequencies (ω).

$$H(\omega) = \sum_{n=0}^{N} h(n) e^{-j\omega n} \quad (3.1)$$

The output, which is the filtered data, is derived from equations 3.2 and 3.3.

$$y(n) = h(n)x(n) \quad (3.2)$$

$$h(n) = y(n) / x(n) \quad (3.3)$$

where h(n) is the filtered coefficient, x(n) is the raw EEG signals and y(n) is the filtered signals. The filtering process is to produce the Delta band, Theta band, Alpha band and Beta band of EEG signals. The plot of the filter coefficient, up to the selected filter length, using specifications stated earlier is shown in Figure 3.4. The calculated filter coefficient will be used to produce filtered data by convolving the filter coefficient with the raw EEG data.

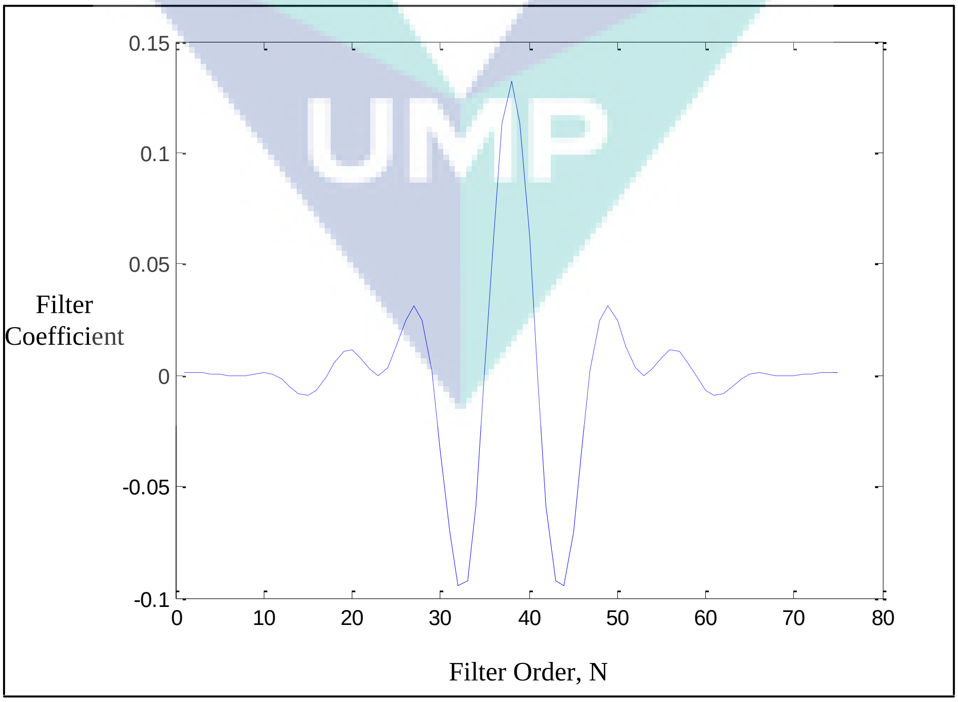


Figure 3.4: A Plot of Filter Coefficient versus Filter Order

3.3 SPECTRAL ANALYSIS OF EEG SIGNALS

After designing the filter to capture the EEG frequency bands, the next step is to perform spectral analysis of the filtered signals using non-parametric technique. In this case, the time-domain EEG signals is converted to frequency-domain signals using Discrete Fourier Transform (DFT). The conversion is implemented using the Fast Fourier Transform (FFT) function in MATLAB. The results will return the DFT value of the filtered EEG data in terms of power spectrum (power spectral density) versus frequency. Equation 3.4 depicts the formula to sample the raw data into finite discrete database on the length of the data, measurement time and sampling frequency.

$$x(t) \Rightarrow x(nTs) \tag{3.4}$$

In equation 3.4, n is the finite length sequence and T_s is the the sampling time. Sampling time is the reciprocal of the sampling frequency. N-point DFT of x(n) is calculated using equation 3.5.

$$X(k) = \sum_{n=0}^{N-1} x(nTs)e^{-j2\pi nk/N} \quad , \quad k = 0,1, \ldots, N\text{-}1 \tag{3.5}$$

In equation 3.5, k represents the harmonic number of the transform component, while N is the length of the sequence. The output of the equation will be in terms of complex component that consist of real and imaginary components. Thus, to secure the real component, which is the power spectral density (PSD), the absolute value of equation 3.5 is calculated, as shown in equation 3.6.

$$PSD = \left|X(k)\right|^2 = \left|\sum_{n=0}^{N-1} x(nTs)e^{-j2\pi nk/N}\right|^2 \tag{3.6}$$

Then, energy spectral density (ESD) is calculated using the results of the power spectrum (PSD) of the filtered data. Here, ESD is measured based on the area under the PSD curve of the specified frequency range of the EEG frequency bands. Following that, the asymmetry energy ratio, relative energy ratio and spectral centroids are calculated, where these parameters are selected as EEG features for this study. Meanwhile, spectral entropy and Z-score technique are used to confirm the stress pattern obtained from the experimental groups.

3.3.1 Power Spectral Density (PSD)

PSD is calculated by performing Fast Fourier Transform (FFT) on the discrete-time data, or sampled data, as stated in equation 3.6. In this study, both the MATLAB's *psd* and *fft* functions were used to compute the PSD value of the filtered data. Here, the *psd* function requires the use of several parameters, such as the number of FFT points, sampling frequency, and size of window with 50% overlapping before applying these parameters to the filtered data. Meanwhile, the *fft* function is applied to the filtered data to act as a comparison to the sprectral analysis using *psd* function. The overlapped window is used to segment the database on the window size and percentage of overlapping. The selected length of the FFT length is 1024. The selection of the length of the FFT must follow the length of the filtered data and the size of the window [7, 61, 99, 112, 177-178]. An example of the PSD plot is shown in Figure 3.5, which depicts the PSD plot for Beta bands (13–30 Hz) where the mean frequency of the band is around 19-20 Hz. The mean frequency of each frequency band is also calculated.

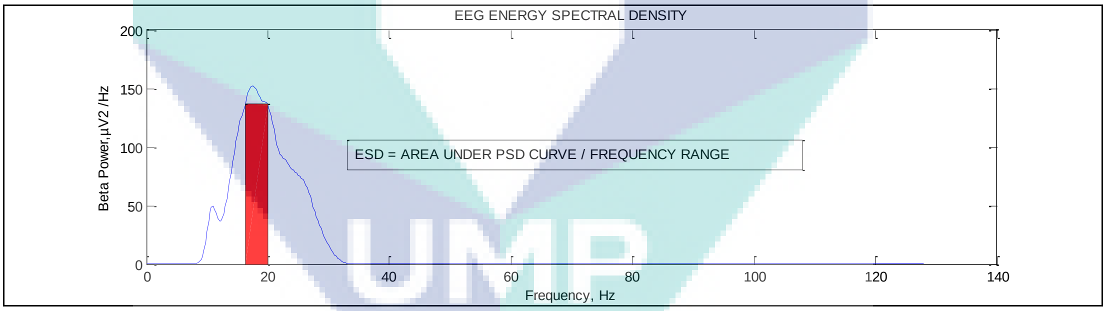


Figure 3.5: Example of a PSD Plot of EEG Beta Band

The plot of the PSD for the overall frequency bands is shown in Figure 3.6. This figure elucidates the characteristics of the EEG sub-bands based on the specified frequency bands from 0.5 Hz to 30 Hz. Beta and Alpha bands are located at the higher frequency but have lower amplitude, while Delta and Theta bands are located at lower frequencies but have higher amplitude. The plot of the PSD characteristic confirms the characteristic of EEG spectrum [7, 8, 16].

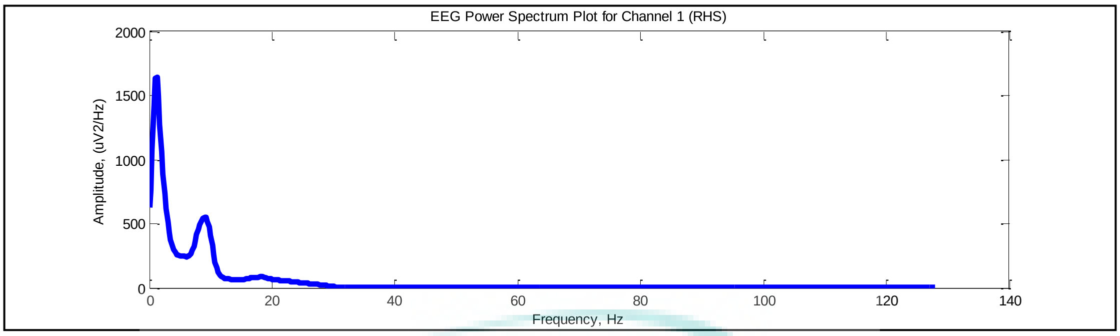


Figure 3.6: Characteristic of PSD of EEG Signals

3.3.2 Energy Spectral Density (ESD)

To determine the features that exist in the EEG signals, the next step is to calculate the values of ESD by performing *trapezoidal* numerical integration on the PSD values over the frequency range using the *trapz* function in MATLAB. This function requires two parameters, for example, *trapz*(X, Y) will compute the integral of Y with respect to X using *trapezoidal* integration. In this study, the value of X is the range of frequency, while the value of Y is PSD, and as depicted by Figure 3.5, the area under the PSD curve is considered as ESD. Since the EEG frequency bands consists of the range of the frequency of EEG signals, ESD is selected as EEG features because it indicates the calculation of the energy of the EEG signals within the EEG frequency bands. In comparison, PSD values indicate the highest power at the mean or peak frequency. In this study, ESD of the EEG frequency bands (δ, θ, α and β) is the mean ESD, namely, the average of the ESD at the right hemisphere and left hemisphere of the human brain as stated in equations 3.7, 3.8, 3.9 and 3.10.

$$ESD(\delta) = \frac{1}{2}\left(ESD(\delta)_R + ESD(\delta)_L\right) \quad (3.7)$$

$$ESD(\theta) = \frac{1}{2}\left(ESD(\theta)_R + ESD(\theta)_L\right) \quad (3.8)$$

$$ESD(\alpha) = \frac{1}{2}\left(ESD(\alpha)_R + ESD(\alpha)_L\right) \quad (3.9)$$

$$ESD(\beta) = \frac{1}{2}\left(ESD(\beta)_R + ESD(\beta)_L\right) \quad (3.10)$$

The reason for taking the average value of ESD for all frequency bands is that the cerebral activities in both hemispheres of the human brain are basically similar, unless the stimulation on the left hemisphere is different from that of the right hemisphere. This can be seen from the EEG datasets at right and left hemisphere, where the values are approximately the same (Refer to Appendix C).

3.3.3 Asymmetry Ratio (AR)

As discussed in the literature review, the main purpose of AR is to show the dominant activity of the brainwaves, either in the right hemisphere or left hemisphere of the human brain. In addition, it is used to compare the neural activities in the left and right hemisphere of the brain. Two basic formulas that can be used to calculate asymmetry are as in equations 3.11 and 3.12.

$$NormalizedAsymmetry \Rightarrow \left(\frac{ESD_R - ESD_L}{ESD_R + ESD_L} \right) \tag{3.11}$$

$$Natural \log Asymmetry \Rightarrow \ln \left(\frac{ESD_R}{ESD_L} \right) \tag{3.12}$$

Equation 3.11 is known as normalized asymmetry where R stands for the power spectrum of EEG frequency bands of the right hemisphere of the human brain and L indicates the power spectrum of EEG frequency bands of the left hemisphere of the human brain, while equation 3.12 is called ln-transformed (natural log) asymmetry using natural logarithm. The expression for normalized asymmetry is bounded between 1 and -1. However, ln-transformed asymmetry is unbounded [28]. Both asymmetry and normalized asymmetry ratios were evaluated in this study.

3.3.4 Relative Energy Ratio (RER)

RER is the ratio of the ESD of each EEG sub-bands over the total energy of the EEG sub-bands. It is described by equations 3.13, 3.14, 3.15, 3.16 and 3.17. RER for each EEG rhythms for all the four groups of subjects involved in the experiment are taken as EEG features and become an input to the classifier. ESD of the EEG frequency bands is determined by calculating the mean of the ESD of the right hemisphere and ESD of the left hemisphere of human brain.

$$TotalEnergy = ESD(\delta) + ESD(\theta) + ESD(\alpha) + ESD(\beta) \tag{3.13}$$

$$RER(\delta) = \left| \log_{10} \frac{(ESD(\delta))}{TotalEnergy} \right| \tag{3.14}$$

$$RER(\theta) = \left| \log_{10} \frac{(ESD(\theta))}{TotalEnergy} \right| \tag{3.15}$$

$$RER(\alpha) = \left| \log_{10} \frac{(ESD(\alpha))}{TotalEnergy} \right| \tag{3.16}$$

$$RER(\beta) = \left| \log_{10} \frac{(ESD(\beta))}{TotalEnergy} \right| \tag{3.17}$$

By computing RER over the EEG frequency bands for four experimental groups, the pattern of stress from the EEG signals might be discovered. Even though the pattern of stress might be clearly shown by RER of Alpha and Beta, this study also considers the stress pattern from the Delta and Theta sub-bands. In order to strengthen the pattern, another EEG feature, the spectral centroids, is applied to the RER of the experimental groups. It is used to search for the dominant EEG energy among the experimental groups.

3.3.5 Spectral Centroids (SC)

Spectral Centroids (SC), one of the features under frequency domain or spectral analysis, can be mathematically expressed as in equation 3.18.

$$C_i = \frac{\sum_{i=1}^{n} F_i |S_i|}{\sum_{i=1}^{n} S_i} \tag{3.18}$$

In this equation, energy is computed from the spectrogram (S) of the RER. F_i is the average frequency weighted by the amplitude of S_i, where i represents the number of samples in the experimental groups. The values of the spectral centroids from each group are used as a target or class for the classification process.

3.3.6 Spectral Entropy (SE)

Another feature that is selected is the spectral entropy (SE), one of the features in the frequency domain. In this study, SE is based on Shannon's entropy and it is chosen to detect the pattern of the stress from the experimental groups by finding the abnormality or uncertainty in the EEG data that might be related to stress. The entropy is mathematically described by equation 3.19. The main purpose of using Shannon's entropy is to indicate which experimental groups might have stress features before proceeding to the classification of the EEG features.

$$SE = -\sum_{i=1}^{n}(RER_i)\log_{10}(RER)_i \qquad (3.19)$$

3.4 EEG DATA RE-GENERATION

In this study, a total of 37 data from 4 experimental groups (180 data; Group 1 – 50 data, Group 2 – 50 data, Group 3 – 40 data and Group 4 – 40 data) were founded corrupted and must be eliminated; 8 EEG data from Group 1, 3 EEG data from Group 2, 13 EEG data from Group 3 and 13 EEG data from Group 4. The corrupted data were identified when there is a high difference between the EEG data of the left brain to that of the right brain. As mentioned in Chapter 2 and section 3.3.2, neural activities in both hemispheres of human brain are normally the same, unless the brain is stimulated by something. Thus, the corrupted data can be easily detected. The corrupted data are then re-created using normally distributed pseudo-random numbers technique. This technique was selected since it can produce positive and negative values. To re-generate the corrupted data, an acceptable noise factor was applied to any good raw EEG datasets in the experimental group. The acceptable noise was determined by keeping the maximum difference between the original data and re-generated data below 10%, which can be achieved by varying the noise factor shown in equation 3.20 and 3.22. Some research had assigned a maximum difference of 5% between the original data and original data with noise [179]. The processes of re-generating EEG datasets by adding noise are described by equations 3.20, 3.21 and 3.22. The noise factor, *F* of 1.5 was used to re-produce an acceptable 37 EEG data. The method to generate oscillation signal to be used as noise using MATLAB function file is described in Chapter 4.

$$Noise = randn(EEGdata) * F \tag{3.20}$$

$$\mathrm{Re}-generatedEEGdata = EEGdata + Noise \tag{3.21}$$

$$F = 1.5 \tag{3.22}$$

3.5 CLASSIFICATION METHOD: k-NEAREST NEIGHBOR (k-NN)

This study uses the k-NN classifier to classify the extracted EEG features. The classifier operates by comparing a testing data (a new sample) with the training data (baseline data). When the *k* neighborhood in the training data is found to match the training data, the classifier assigns the testing data with the class that appear more frequently in the neighborhood of *k*, which indicates the number of nearest neighbors in the classification. In order to determine the class that match the testing data to the training data, the value of *k* must be varied. In this research, the values of *k* are changed in accordance with the size of the training data. Typically, a small value of *k* is required for the classification with a maximum value of 10 [1, 7, 33, 94, 112, 180-181], and the default value of *k* is 1. For example, if *k* = 1, the test data will be assigned to the class of its nearest neighbors, where *k* must be a positive integer. The partition of the dataset for training and testing in k-NN classification process can be done according to 50:50 ratio, 60:40 ratio, 70:30 ratio and 80:20 ratio. In this study, two different training-to-testing ratio, which are 50:50 and 70:30, were used to evaluate the EEG datasets in order to determine the ratio that would produce good classification accuracy.

Besides the k-value and the ratio of partitioning data, other parameters that must be taken into account in classification are distance and rule. In order to identify neighbors, the distance and rule of k-NN classifier must be chosen. There are five types of k-NN distance (Euclidean, City block, Cosine, Correlation and Hamming) and three k-NN rule (Nearest, Random and Consensus). The k-NN distance in term of Euclidean, City block and Cosine are described by equation 3.24 to equation 3.26 respectively. The default neighborhood setting for distance and rule is Euclidean and Nearest, respectively, and are the commonly used settings by reseachers [1, 33, 94]. Hamming distance is not applicable in this study since it more suitable for classifying binary data. In this study, to search for the object similarity in the k-neighborhood, all k-NN rules are employed in the classification in conjunction with Euclidean, City block and Cosine of k-NN distance, as shown in equations 3.23, 3.24 and 3.25, respectively. The method to implement k-NN classification using MATLAB will be elucidated in Chapter 4.

Equation 3.23 describes the Euclidean distance in k-NN classification. X_i or X_j is either the training or the testing data, where *i* and *j* indicate the index of the data, while *k* is the counter for the length of the training data (*n*).

$$d(X_i, X_j) = \sqrt{\sum_{k=1}^{n} (X_{ik} - X_{jk})^2} \tag{3.23}$$

Equation 3.24 describes the city block distance in k-NN classification. X_i or X_j is either the training and the testing data, where *i* and *j* indicate the index of the data, while *k* is data counter. The distance is the summation of the difference between the data and the results that was assigned to the class that came out more frequently in the neighborhood of *k*.

$$d(X_i, X_j) = \sum_{k=1}^{n} | X_{ik} - X_{jk} | \tag{3.24}$$

Cosine distance is elaborated by equation 3.25. Similarly, X_i or X_j is either the training and the testing data, *k* is the data counter and *n* is the length of the training data. Cosine distance involves the summation of the multiplication of the data over the square root of the data that was assigned to the class that came out more frequently in the neighborhood of k.

$$d(X_i, X_j) = 1 - \sum_{k=1}^{n} \frac{(X_{ik} * X_{jk})}{\sqrt{| X_{ik} |^2} * \sqrt{| X_{jk} |^2}} \tag{3.25}$$

Each of the three k-NN rules (nearest, random and consensus) that can be used to classify EEG dataset works differently. The nearest rule applies the majority rule with nearest point, the random rule utilizes the majority rule with random points (use random tie-breaker), while the consensus rule applies the opposite that of the majority rule. However, by default, the nearest rule is used in k-NN classification, where the test data are assigned to the class that consists of the majority of the k-nearest neighbors. For consensus rule, testing data that consists of k-nearest neighbors from different class are not assigned to the class but the NaN will be assigned to the output class. In this study, the k-NN classifier was evaluated using Euclidean, City block and Cosine distance, while all rules were used in order to find the best classification results.

3.6 CLASSIFICATION PERFORMANCE: ACCURACY, SENSITIVITY AND SPECIFICITY

In order to produce the best classification performance, the classifier must be tested for accuracy, sensitivity and specificity [158, 168, 170, 180]. Accuracy is defined as the closeness of the measurement to its true value, sensitivity is described as the true positive that is correctly identified, and specificity indicates the true negative that is correctly identified. Accuracy, sensitivity and specificity can be calculated using equations 3.26 3.27 and 3.28, respectively. In order to make proper computation of accuracy, sensitivity and specificity in the classification process, a confusion matrix need to be built.

$$Accuracy = \frac{(TP + TN)}{TP + TN + FP + FN} \tag{3.26}$$

$$Sensitivity = \frac{(TP)}{TP + FN} \tag{3.27}$$

$$Specificity = \frac{(TN)}{TN + FP} \tag{3.28}$$

The confusion matrix for the 2 x 2 matrices of the datasets is shown in Table 3.4.

Table 3.4:
Confusion Matrix for 2x2 Matrices [182-184]

	Target Class	
Actual Class	TP	FP
Actual Class	FN	TN

In the matrix, TP represents True Positive, FP is False Positive, FN indicates False Negative and TN for True Negative.

However, this study involved the feature vectors of 4 x 4 matrices, which are 4 groups with 4 EEG frequency bands. Therefore, the confusion matrix for 2 x 2 matrices are modified to produce confusion matrix for 4 x 4 matrices as shown in Table 3.5. The reference for the 4 x 4 confusion matrix can be obtained from the website under the topic of evaluating classification model [169, 182-184].

Table 3.5:
Confusion Matrix for 4x4 Matrices [182-184]

		Predicted Class			
		A	**B**	**C**	**D**
Actual Class (Known data)	**A**	TP_A	eAB	eAC	eAD
	B	eBA	TP_B	eBC	eBD
	C	eCA	eCB	TP_C	eCD
	D	eDA	eDB	eDC	TP_D

Based on Table 3.5, accuracy, sensitivity and specificity can be calculated using equations 3.29, 3.30, 3.31 and 3.32.

$$Accuracy = \frac{(TP_A + TP_B + TP_C + TP_D)}{TP_A + TP_B + TP_C + TP_D + error} \tag{3.29}$$

$$error = eAB + eAC + eAD + eBA + eBC + eBD + eCA + eCB + eCD + eDA + eDB + eDC \tag{3.30}$$

$$Sensitivity_A = \frac{(TP_A)}{(TP_A + eAB + eAC + eAD)} \tag{3.31}$$

$$Specificity_A = \frac{(TP_B + TP_C + TP_D)}{(TP_B + TP_C + TP_D + eBA + eCA + eDA)} \tag{3.32}$$

$Sensitivity_A$ is for sensitivity of Class A, while $Specificity_A$ indicates the specificity of Class A. The overall results of accuracy, sensitivity and specificity obtained from the experiments are detailed in Chapter 5.

3.7 CLASSIFICATION ERROR

Besides sensitivity and specificity, the classification performance is also measured based Mean-Square Error (MSE) when varying the distance of k-NN classifier. MSE is calculated from k = 1 to k = 10 when the distance of the k-NN are changed from Euclidean to City block and Cosine, respectively.

3.8 VALIDATION OF k-NN CLASSIFIER

The performances of the classifiers are evaluated through cross-validation, where the accuracy of the classifier is re-calculated by implementing multiple partitions of the datasets that was used for training and testing. The accuracy obtained from the validation process will be compared with that acquired from normal classification to determine whether or not the classifier is robust. Generally, the classification cross-validation can be achieved using k-fold cross-validation and Leave-One-Out (LOO) cross-validation. In k-fold cross-validation, k is the number of fold employed in the validation; if a 5-fold cross-validation is applied, 4 fold are used for training while one is for testing. This means that, in k-fold cross validation, all data in the datasets are used for training and testing. However, the size of the datasets will determine the number of fold to be employed in the cross-validation, the bigger the number of data in the dataset, the bigger is value of *k* [168, 170, 185-189, 201]. In this study, a 10-fold cross-validation is chosen as the size of dataset matrix is 180x4, which is also the most used cross-validation to validate k-NN classifier performance [168, 170, 185-189, 201].

The k-NN classifier is also validated using LOO technique, in which case one data in the datasets is selected for testing while the remaining data becomes the training data. The process begins by choosing data number one as the test data and making the rest of the data as the training data. Then data number two is selected as the test data, while the remaing data becomes the training data. This process continues until the counter reaches the last data in the datasets.

In both classifier cross-validation techniques, the accuracy obtained from each experiment is then used to produce average accuracy, where the results of the validation are shown and discussed in Chapter 5.

3.9 CLUSTERING METHODS: FUZZY C-MEAN and FUZZY K-MEAN

In order to support the classification results by k-NN classifier to generate stress index based on quantitative method, the extracted features from the EEG signals of the experimental groups are clustered using both FCM and FKM. The fuzzy technology is widely used by researchers, especially in identifying human emotions. FCM and FKM are unsupervised learning algorithms employed to compute the classification accuracy based on the membership degree of the cluster center and the centroids of the cluster center respectively [160-161, 164, 166, 189, 201]. Both FCM and FKM depend on the number of clusters used to partition the datasets, hence, in this study, the datasets are partitioned into four clusters as datasets from four groups are involved in the experiments.

The first step in the FCM process is to define the required number of clusters in order to partition the datasets, which is followed with the calculation of the cluster center. The next step is to calculate the degree of membership, and finally, is the calculation of the clustering accuracy [161, 190]. The outputs from FCM are cluster center matrix, membership function matrix and the objective function. Membership is graded from 0 to 1, and four clusters will be produced as there are four experimental groups in the datasets. The results from the membership are then used to index the clustered datasets. The datasets are then plotted to determine the dispersion of the datasets in the neighborhood of the cluster center. The objective function of FCM (J) is calculated using equation 3.33.

$$J = \sum_{k=1}^{n}\sum_{i=1}^{c}\mu_{ik}^{m} \| x_k - v_i \|^2 \tag{3.33}$$

In the equation, n is the total number of data points, c is the number of clusters, x_k is the kth data point, v_i is the cluster center, and μ_{ik} is the degree of membership of kth datapoint and ith cluster center, where the value ranges from 0 to 1 with weighting effect of m. Meanwhile, $\|x_k\text{-}v_i\|$ is the difference between datapoints and its cluster center. Membership degree can be determined using equation 3.34.

$$\mu_{ik} = \frac{1}{\sum_{j=1}^{c}\left(\left(\frac{\| x_k - v_i \|}{\| x_k - v_j \|}\right)\right)^{2/(m-1)}} \tag{3.34}$$

which indicates the degree the datasets belong to the cluster center of v [165, 190-191]. The index of the membership matrix is used as a class to k-NN classifier to ascertain the accuracy of the FCM clustering. To calculate the accuracy of the clustering, the clustered datasets is then used as an input data to k-NN classifier.

In FKM, which can be implemented using equation 3.35, the number of clusters is used to determine the location and the index of the centroid in the datasets, and consequently, each cluster has k centroids. Squared Euclidean distance is employed to determine the centroids of the data in each cluster, and the data points are then clustered based on the minimum Euclidean distances. Finally, the new center of the cluster is calculated based on the average of the data in the cluster [7, 112, 164, 192]. The new location index is then utilised to determine the new data cluster, which will then be used as an input to k-NN classifier.

$$J = \sum_{k=1}^{n}\sum_{i=1}^{c}\| x_k - v_i \|^2 \tag{3.35}$$

In Equation 3.35, n is the total number of data points, c is the number of clusters, x_k is the data points, v_i is the cluster center, while $||x_k\text{-}v_i||$ represents the Euclidean distance between kth data points x_k and the cluster center, v_i. J is the objective function of FKM where the main difference between FCM and FKM is the use of membership degree, μ.

3.10 STATISTICAL ANALYSES

This section will describe the theory of the statistical techniques to confirm the stress index assignment using Z-score technique, normality test, ANOVA and Pearson correlation study (scatterplot). Here, the Z-score is used to verify the stress index. Meanwhile, the normality test is applied to test the normality of the selected features before classification process. Next, ANOVA is selected to check the significance of the selected features. Finally, Pearson correlation study is performed to determine the correlation among the selected features when there is a change in brain cognitive states.

3.10.1 Z-score Technique

Z-score technique is applied in this research in order to confirm the stress pattern obtained from the energy ratio, spectral centroids and entropy. Z-score is used to calculate the mean of each group and their standard deviation, and is expressed as in equation 3.36.

$$Z - score = \left({(x - mean(x))} \big/ {\sigma(x)} \right) \tag{3.36}$$

The equation will produce a value of less than -1, zero and greater than 1. The results from Z-score will be used to validate the index assignment based on energy ratio, spectral centroids and classification results. Based on the formula of Z-score, it can be used to generate 3 indices from the scores (less than -1, zero and greather than 1). Hence, this technique is practicable and suitable to verify the assignment of stress index based on the results of the relative energy ratio and classification of EEG features.

3.10.2 Normality Test

The normality test is performed in order to determine the Gaussian pattern of each EEG dataset. The test is implemented for the filtered data, namely, after obtaining the energy or power spectrum of the data, and is also performed after the corrupted data is re-generated. The normality test is carried out using the histogram plot in MATLAB and the results of the test are discussed in Chapter 5. The testing for normality is important in order to enhance the classification performance.

3.10.3 ANOVA

In this study, ANOVA is used to indicate the significance of the extracted features, relative energy ratio and spectral centroid from different sets of data, where the significance of the features of different groups is based on the mean analysis of the data. In ANOVA, the dependent variables and independent variables need to be defined [173, 193]. In this study, ANOVA is used is to determine the degree of stress of the four experimental groups based on the change of Alpha and Beta power across the group. The cognitive states of Group 1, Group 3 and Group 4 is eyes-closed, while that of Group 2 is eyes-open, where members of Group 2 are requested to perform an IQ test. The dependent variables are the groups and the independent variables are the extracted features from the groups. Even though four experimental groups are involved in the study, since the main focus is to determine the significant difference between the members of Group 1 (who are asked to be in a relaxed state) and that of Group 2 (who had to undergo the IQ test), ANOVA is used instead of MANOVA. Furthermore, ANOVA will indicate the significant difference within the group and between groups by producing the output value of F and *p*. F-value indicate the degree of freedom, while *p*-value represents the degree of confidence [173]. The large value of F indicates the high variability between the groups. Meanwhile, the small value of F indicates the low variability between the groups. The *p*-value of 0.05 or less will indicate the significant difference on the dependent variable in the groups. Therefore, the extracted features from the two groups are said to have a significant difference if the p-value less than 0.05. In order to produce a good result from ANOVA analysis, the data for each variable must be normally distributed [173].

3.10.4 Pearson Correlation Study

Beside ANOVA, in order to discover the linear relationship between the change of the cognitive states (eyes-closed and eyes-open) with the change of the EEG power spectrum, Pearson correlation (scatterplot) study is conducted. The correlation results between the cognitive state and its EEG power spectrum will be indicated by the Pearson correlation coefficient, *r* and the significant value, *p*. The range of the correlation coefficient is between -1.00 to 1.00 which indicates the correlation can be positive and negative where 0 will indicate zero correlation between the two variables being correlated. The negative and positive correlation indicate the direction of the relationship. Regardless of the positive and negative correlation, the correlation coefficient will indicate the strength of the correlation results where the strength can be categorized into 3 categories; small correlation strength when *r* between 0.10 to 0.29, medium correlation strength when *r* between 0.30 to 0.49 and large or strong correlation strength when *r* between 0.50 to1.0. Meanwhile, the variables are significant if the significant value of the linear relationship less than 0.05 ($p < 0.05$). The results of the statistical analyses (ANOVA and Pearson correlation coefficient) of the extracted EEG features are discussed in Chapter 5.

CHAPTER FOUR

METHODOLOGY

4.1 INTRODUCTION

The description of how stress index is generated starting from the raw EEG datasets to the classification of the stress features is detailed in this chapter. Section 4.2 describes the flow chart of the experimental steps conducted in this study. Data collection, data re-generation, equipment set-up and validation are elaborated in Section 4.3. The EEG measurement protocols are detailed in Section 4.4. Meanwhile, the pre-processing of the EEG datasets, including the removal of artifacts, filter design and normality test, is discussed in Section 4.5. The major portion of this chapter is covered by Sections 4.6 and 4.7 which, respectively, describe the spectral analysis of EEG signals and the selection of EEG features that might relate to stress. This is followed with discussion on the procedures to classify and cluster the extracted stress features in Section 4.8 and Section 4.10, while Section 4.9 describes the procedures on how to validate performance of the classifier. Then, the statistical analysis of the stress features is presented in Section 4.11. The assignment of the stress index based on the selected EEG features and the classification results is explained in Section 4.12, and finally, Section 4.13 describes the steps taken to verify the stress index.

4.2 RESEARCH PROCESS FLOW

This study involves several process flows as illustrated by Figure 4.1, Figure 4.2 and Figure 4.3. Here, Figure 4.1 illustrates the main process flow of the research. According to Figure 4.1, it is obvious that the vital parts in process flow which are also the milestones in this research are the identification and selection of the EEG features that might relate to stress, classification of the features, construction of the stress index based on the selected features and the verification of the stress index.

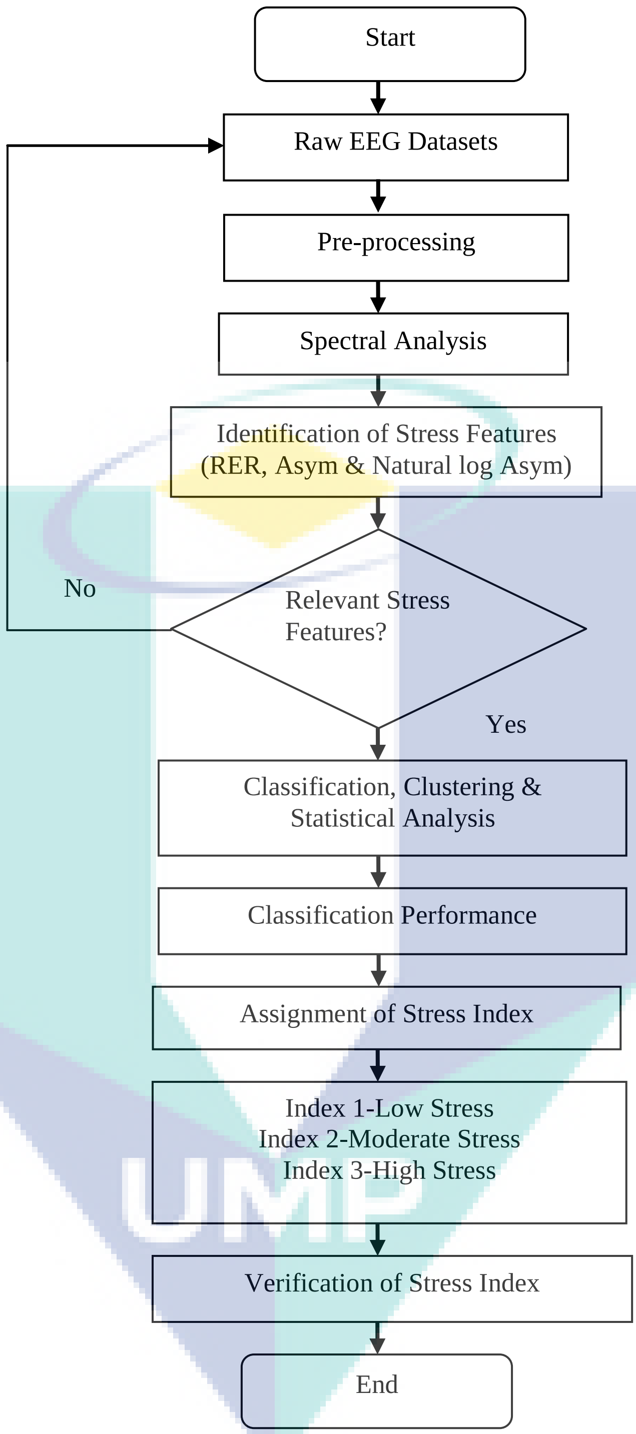


Figure 4.1: Overview of Research Process Flow

Looking at the main process flow (Figure 4.1), the research employs the available raw EEG datasets obtained from the four groups of subjects involved in the experiment. The data is then pre-processed to remove the artifacts and filtered at the required EEG frequency bands by applying band-pass filter with the specified cut-off frequencies to the noise-free data. Next, the non-parametric or spectral analysis is performed to identify the EEG features that might relate to stress. If the features are related to stress, the features will be tested for normality first, before selecting the features as an input to the classifier for the classification process as labeled in the process flow as relevant stress features, otherwise, a new raw EEG data must be selected and the process repeated. As discussed in Section 2.5 in Chapter 2, there are various features in non-parametric analysis methods that can be selected and analyzed to idenfity the existence of stress in human. Among them, 3 features were selected and tested for stress identification process; Relative Energy Ratio (RER), Asymmetry and Natural Log Asymmetry since these features can detect the changes in human cognitive activities when stress stimuli are applied to human [56-61, 72, 75-77]. Next, the selected features are fed to classifiers. The selection of the features in determining stress index will be based on the classification performance. In this thesis, k-NN is selected as a classifier due to its robustness and less complicated supervised learning algorithm, while FCM and FKM are used to cluster the features extracted from EEG signals. These clustering techniques are employed to support the classification results by k-NN classifier. The performance of the classification is measured based on accuracy, sensitivity and specificity.

The robustness of the classification is tested by applying k-fold cross-validation. To determine whether or not the selected features are significant, statistical analysis is conducted on the features using ANOVA. Beside ANOVA, the study uses Pearson correlation in order to know the linear relationship between the variables being studied with the change of cognitive states. Once the results of the classification are obtained, the index that represents the level of the stress is assigned. The indices are verified using Z-score technique and the range of the EEG data from each of the four groups involved in the experiment. In addition, every EEG data obtained from the four groups are tested whether or not they belong to the correct index. The stress index is developed from the average value of the relative energy ratio, spectral centroids and entropy of EEG Alpha and Beta power spectrum from each group. Hence, it is vital to verify each EEG data in each group and their index. The process flow for the pre-processing and spectral analysis is shown in Figure 4.2 and Figure 4.3, respectively.

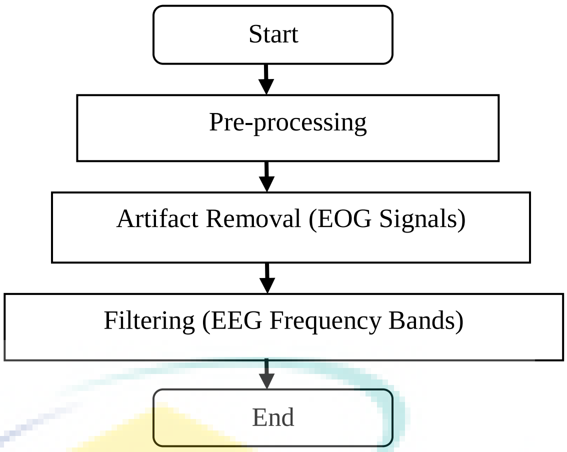


Figure 4.2: Process Flow of EEG Pre-processing

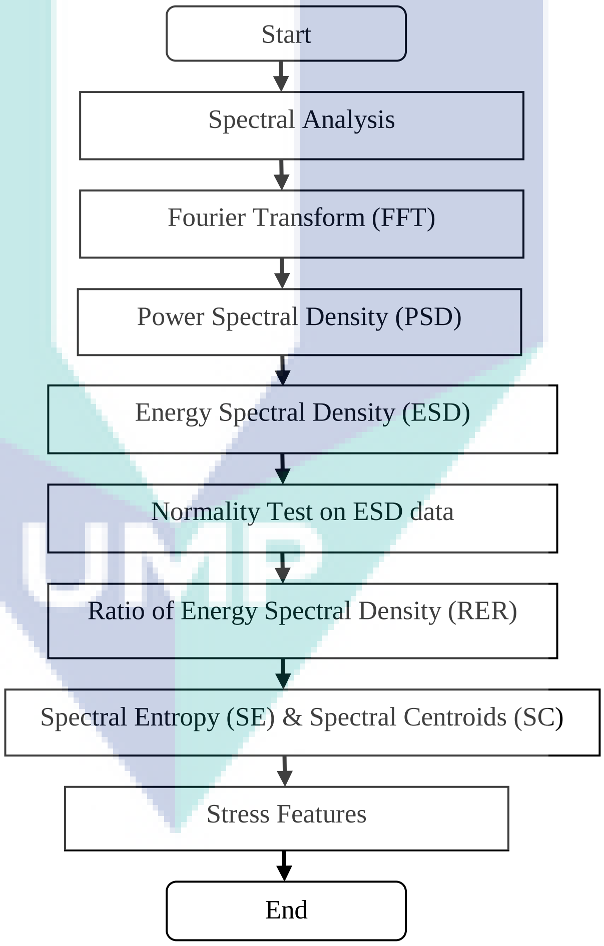


Figure 4.3: Process Flow to Identify Stress Features

Referring to Figure 4.2, the major artifacts or noises in EEG measurements come from EOG signals. Hence, the artifacts in EEG signal are removed by setting the threshold value in the algorithm, where the procedure for removing the artifact has been detailed in Section 4.5. Once the EEG datasets are free of artifact, the data are band-pass filtered to capture the required frequency. In this research, datasets with frequency ranging from 0.5 Hz to 30 Hz are selected, which comprise of Delta band (0.5-4 Hz), Theta band (4-8 Hz), Alpha band (8-13Hz) and Beta band (13-30Hz). Subsequently, the filtered EEG datasets are converted to power spectrum density using Fourier Transform technique, and then tested for normality using Excel and MATLAB. The procedure has been detailed in Section 4.5.3.

This study implements spectral analysis which involve processes as illustrated in Figure 4.3 in order to recognize the EEG features that can be used to detect stress. As discussed in Chapter 2, this technique is one of the non-parametric analysis technique which is widely used by researchers to extract features from non-stationary data such as EEG signals. According to Figure 4.3, the Discrete Fourier Transform (DFT) is applied to the filtered EEG datasets. For the fastest computer computation, the signal processing uses Fast Fourier Transform or FFT [7, 112, 177-178]. As a result, the time-based data platform is converted to frequency-based data platform or power spectrum. The reflection of the power at each frequency by power spectrum is called PSD, having unit of $\mu V^2/Hz$. In this study, ESD, obtained by dividing the area of the PSD curve with the range of the specified frequency, is considered as one of the feature of EEG sgnals. ESD is selected as a feature, instead of PSD, because it covers the overall energy distribution for each frequency band, while PSD selects the highest energy at the peak frequency. Thus, the value of ESD is relatively smaller and more precise than PSD as it covers the energy in the entire frequencies of EEG signals. The stress features are selected by calculating the ratio of ESD at each frequency band over the total energy of overall band. The ratio of ESD from each EEG rhythms are counted. However, since the stress pattern is indicated by the changes in the energy of Beta and Alpha bands, the focus is only on the ESD of Beta and Alpha bands.

Beside ESD, SC is selected as stress feature as well since the feature were used by researchers to search for the dominant energy or frequency in a certain set of data being studied [118, 128, 130-134]. The feature is applied to the four group of EEG datasets for each frequency band, while entropy is used to detect the stress pattern from the four groups of EEG datasets. Hence, in this study, the ESD ratio or RER and SC are utilised as the stress features for the classification process. In the classification process, RER are used as input features for the training and testing, while SC are used as classes. The identification of the stress features by the classifier can be effectively implemented using the spectral analysis technique [146, 150, 168, 170].

4.3 DATA COLLECTION AND RE-GENERATION

This section discusses the process on how to select the data, set-up and buy-off the equipment. Equipment validation must be done prior to EEG measurement.

4.3.1 Selection of Datasets

The study began by collecting the EEG data from several experiments. 180 EEG datasets were collected from 140 non-smoking healthy subjects whose age range from 20 to 50 years old. The datasets were categorized into four groups. Group 1 represents Eyes-Closed (EC) state consisting of 50 EEG data taken from 50 subjects after they have answered the same psychoanalysis questionnaires. Group 2 represents Eyes-Open (EO) state consisting 50 EEG data obtained from 50 subjects after they had performed IQ test. Meanwhile, Group 3 and Group 4 represent EC state, each consisting of 40 EEG data taken from 40 subjects before and after the subjects performed Horizontal Rotation (HR), respectively.

In the process of the collecting EEG data, of all the EEG data collected, a total of 37 data were found to be corrupted (Group 1 – 8 data; Group 2 – 3 data; Group 3 – 13 data; Group 4 – 13 data), and the data were substituted with new data produced by implementing normally distributed pseudo-random number technique. This technique will generate oscillation data (negative and positive data). In this case, to produce the new data, an acceptable noise factor was applied to any good raw data in each group. In terms of maximum value and standard deviation, the divergence between the re-generated data and the original data were maintained below 10%. The noise factor of 1.5 was employed to regenerate the 37 EEG data, which produced data that have characteristics that are almost similar to that of the removed data.

4.3.2 Equipment Set-up

The g.MOBIlab with Bipolar EEG gold-plated and silver-plated electrodes were used in this research for recording EEG signals from each subject. The impedance of the electrodes was measured using z-checker and found to be below 5 kΩ [7, 16, 194]. The impedance of the electrodes are vital so as not to introduce additional noise to the EEG reader. In this study, two electrodes were placed on the subject's prefrontal area of human brain region and are denoted by Fp1 (left channel) and Fp2 (right channel). One electrode was placed at the center of the forehead to represent the ground, which is denoted by Fpz. Meanwhile, the reference electrodes were placed at the earlobes, which are denoted by A_1 and A_2. The placing of the electrodes are in accordance with the International 10-20 System. Prior to placing the electrodes on the forehead and earlobes, the electrode cap is filled with gel to remove noise originating from the skin. Figure 4.4 illustrates the set-up of the EEG instrument.

The EEG signals from the electrodes are sampled at 256 Hz (3.9 milliseconds). The condition of the electrodes was verified by conducting the measurement for a few seconds without placing the electrodes onto the subjects. When drift or abnormal reading is observed, the condition of the electrodes need to be checked to exclude the noise from being included in the EEG measurement. If the abnormal reading keeps occurring, then the EEG electrodes need to be replaced with a new one. This research identified that the artifacts or noises from measuring instrument could be eliminated by either keeping electrode impedance below 5 kΩ [7, 16] or changing the faulty electrode with a new one. Some of the EEG electrodes that have been frequently used might be faulty due to wear and tear. Therefore, before starting any experiment, the condition of the EEG electrodes must be checked and only those in good condition are used in the experiment.

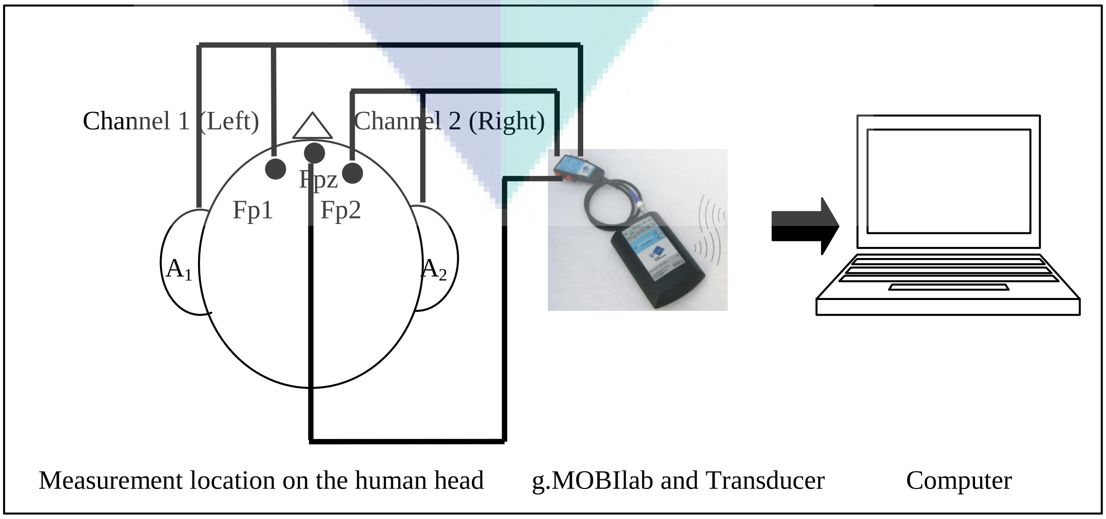


Figure 4.4: EEG Equipment Set-up

In this study, two channels (right and left) are used to capture the EEG signals from the subject's forehead, ares which area are associated with human emotions. Previous studies have found that the minimal number of electrodes is sufficient to capture the EEG signals for determining human trait or behaviors [195-196, 202-203]. After measuring EEG signals, the signals is then sent to the computer using wireless method. In order to receive the EEG signals from the EEG amplifier, the MATLAB SIMULINK block diagram is constructed as illustrated in Figure 4.5.

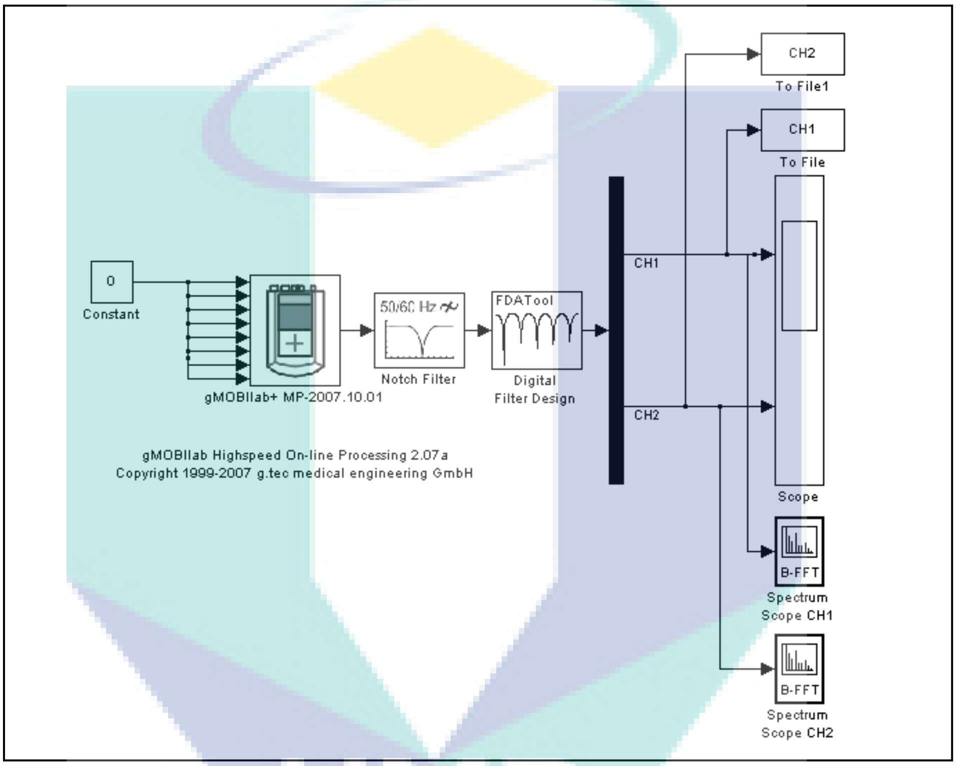


Figure 4.5: SIMULINK Block Diagram for Capturing EEG Data from EEG Amplifier

A notch filter is used to eliminate noise from the power line and a digital filter is used to filter the EEG signals. Here, the frequency of the notch filter is set to 50 Hz. It indicates that the filter will reject any signal received from the EEG amplifier with frequency of 50 Hz and above. Following this, a de-multiplexer is used to extract data from the raw EEG signals into Channel 1 (Right) and Channel 2 (Left), which are then saved in MATLAB m-file for analysis. Even though the transmitted EEG signals have gone through notch filter and digital filter design, the EEG signals for the Channel 1 and Channel 2 is still considered raw since the signals have yet to be filtered from its main interference signal which come from eyes-movement and eyes-blink. Besides, the transmitted EEG signals have yet to be filtered to their sub-bands. Therefore, the

MATLAB coding is developed to obtain the real EEG data and then, to reveal the EEG features from the captured EEG data.

4.3.3 Equipment Validation

The equipment was buy-off in 2 areas, the EEG reading in the specified frequency range (0.5 – 50 Hz) and the EEG reading from 2 electrodes (left and right). 90 experiments were carried out to test the equipment reliability. The input sinusoidal signal from a signal generator was applied to the electrodes of the g-MOBIlab and the output signal was displayed on the oscilloscope. The input signal was set at 10 different frequencies that covers the range of EEG frequency: 0.5 Hz, 1 Hz, 2 Hz, 5 Hz, 10 Hz, 15 Hz, 20 Hz, 30 Hz, 40 Hz and 50 Hz (refer to Appendix D). The setting of frequency of the input signal was done to follow the characteristics of the EEG signals in term of frequency. Meanwhile, the input voltage was varied from 100 μV to 500 μV. In terms of amplitude, there is an offset of 0.96 μV between Channel 1 and Channel 2, where the reading taken by Channel 2 is 0.96 μV higher than that of Channel 1 (refer to Appendix D). However, no frequency difference is detected in both channels. The results of the equipment buy-off show that the EEG reading measured by g-MOBIlab is reliable and the equipment can be used for EEG measurement.

4.4 EEG MEASUREMENT PROTOCOL

In this research, in order to generate two conditions of the cognitive state (EC and EO), the subjects were involved in two activities; first, they were asked to be in a relax state and not do anything, and second, they were requested to answer IQ test questions. For the first activity, in a controlled room, the subjects are instructed to idly sit in a chair with their eyes closed and covered with eye mask while EEG measurements are being taken. This condition is called EC state, and the recording period for this activity is three minutes. However, before commencing the measurement, the system is initialized for five to ten seconds to check the waveforms for any errors or abnormalities. The measurement will only start when no error is detected, otherwise, it will not carried out as to ensure that instrumentation noise is not included in the EEG measurement. The condition of the EEG amplifier, transducer and electrodes are also thoroughly checked for any discrepancies.

For the second activity, to stimulate the EO cognitive state, subjects are required to answer 20 IQ test questions within ten minutes, while their brain activities are simultaneously captured. In order to reduce noise that might be added-up to the EEG data, the subjects are requested to minimize their movements while answering the questions. Offline data processing on the captured signals are performed using advanced signal processing algorithm developed in SIMULINK and MATLAB. The questions in IQ test are created based on the modified Ravens' Standard Progress Matrices (SPM) that focuses only on visual and logical thinking. The Graphical User Interface (GUI) is created to ease the IQ assessment. The EEG measurement protocol is stated in Table 4.1.

Table 4.1:
EEG Measurement Protocol

5-10 seconds	3 minutes	Within 10 minutes
System initialization & checking waveforms condition	EEG recording for EC state (do nothing); Group 1, 3 & 4	EEG recording for EO state (IQ test); Group 2

4.5 PRE-PROCESSING OF EEG DATASETS

This section describes the procedures to remove artifacts or noises from EEG data, to design an appropriate band-pass filter to obtain the Delta, Theta, Alpha and Beta frequency bands from the EEG signals and the energy of each frequency band. Also, in this section is the explanation of the steps required to test the EEG data in terms of Energy Spectral Density (ESD) for any abnormality. The normality test is conducted to ensure the ESD of the EEG signals has the normal Gaussian distribution.

4.5.1 Artifact Removal

The main artifact in this study comes from eye-movements or eye-blinks, which are known as ocular artifact or EOG. Therefore, the EOG signals must be separated from the EEG signals as the characteristics of EOG signals are almost similar to that of EEG signals.

In this study, the EOG signals are removed from the time-based raw EEG signals by setting the threshold values of ± 100 µV inside the signal processing algorithm. By setting this threshold, the program will automatically remove signals that have amplitude above or below ± 100 µV, regardless of the frequency, and consequently, the program might change the length of the EEG data. Other artifacts, such as Electromyogram (EMG), Electrocardiogram (ECG), power line and sweat, give minimal interference to the EEG signal. EMG has amplitude and frequency much higher than EEG signals, while ECG has frequency below 1 Hz, and the noise originating from the power line has a frequency above 50 Hz. The noise from the power line can be removed by applying notch filter [197].

4.5.2 Frequency Filtering

The research utilizes Finite Impulse Response (FIR) band-pass filter to capture the required frequencies from the raw EEG signals as it is very stable, easy to use and not requiring a lot of filter parameters [7, 16, 60-61, 177-178]. The filter, applied after the process of artifact removal, produces signals in the range of 0.5 Hz to 30 Hz that cover the Delta, δ (0.5-4 Hz), Theta, θ (4-8 Hz), Alpha, α (8-13 Hz) and Beta, β (13-30 Hz) of the EEG sub-bands. The right coefficients of the filter are determined as shown in Figure 4.6, in which the filter order is varied until the right coefficients are obtained. The process of finding suitable filter order is implemented by checking the curve at the high-pass frequency and low-pass frequency edges of the frequency response of the filter that meets the cutoff specifications as stated in Table 3.1. Once the correct filter coefficients are obtained, the EEG signals is filtered by applying the coefficients before performing Fourier Transform to the filtered signals in order to convert the signals from time-domain to frequency-domain. This study applies window-based filter on the clean or noise-free EEG data, where the combination of filter length or filter order with window-based filter produces the desired filter coefficients to filter those data. As elucidated by Figure 4.6, the filter coefficients depend on the results of the filter frequency response produced from the calculation of the filter specifications in term of filter order, filter type, cut-off frequencies and type of window. Here, cut-off frequencies are required to produce frequency bands of EEG signals. Meanwhile, the length of the filter is used to tune the frequency response of the filter to meet the specified cut-off frequencies (refer to Appendix E).

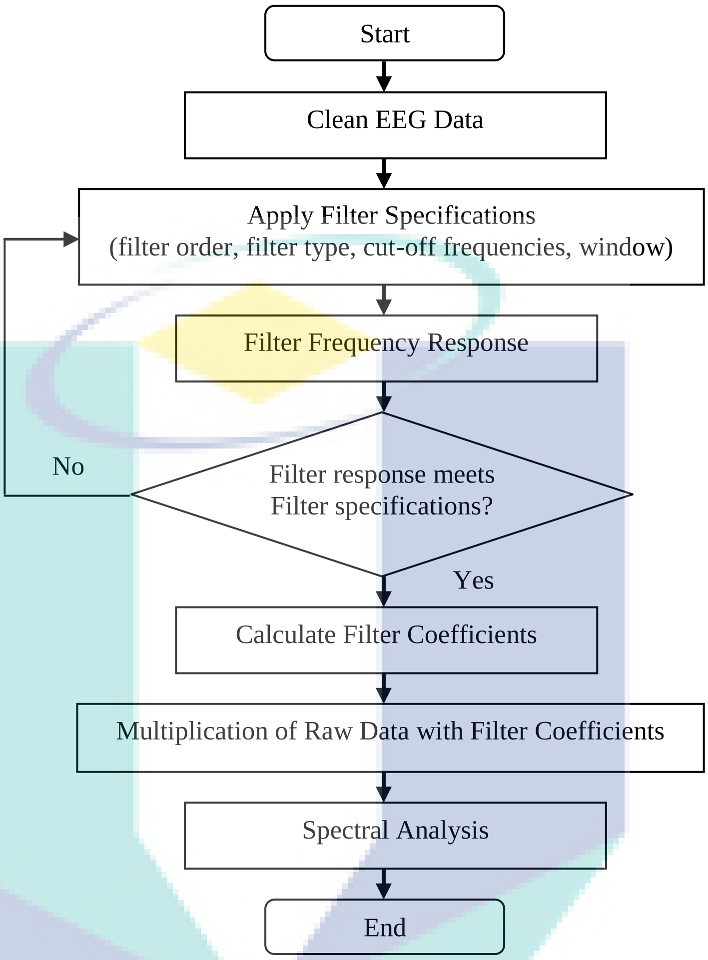


Figure 4.6: Flow Chart for the Determination of Filter Coefficients

In MATLAB, the function file *fir1* is used to generate the filter coefficients, and the filter coefficients are then multiplied by the clean EEG raw data using MATLAB function file *conv*. The filter is designed by adjusting the filter order so that the curve or 'roll-off' area of the filter's frequency response justifies the cut-off frequency. The important parameters after implementing the filter specifications are the frequency response and coefficients of the filter and the results obtained from MATLAB will validate whether or not the filter meets the filter specifications. The process of this research will not proceed to the spectral analysis stage unless the filter is properly designed to produce the frequency bands of the noise-free EEG signals.

4.5.3 Test of Normality

This section explains the detail process of the normality test on the EEG data. In this study, the data is tested for normality using histogram plot after the data converted to the PSD (frequency-based) as illustrated by Figure 4.7. The power spectrum for all frequency bands and groups present in the EEG data will be checked for mean and standard deviation using MATLAB. The shape of the graph should follow the normal Gaussian distribution (bell-shaped), otherwise, the experiment must be repeated from the beginning as shown in Figure 4.7. The MATLAB function file, *psd* is used to convert the clean EEG signal to power spectrum using *Hamming* window with 50 % overlap. Compared with time-domain, the Gaussian distribution of the EEG data can be clearly seen in the frequency-domain, and more EEG features can be extracted when the signals are analysed in frequency domain.

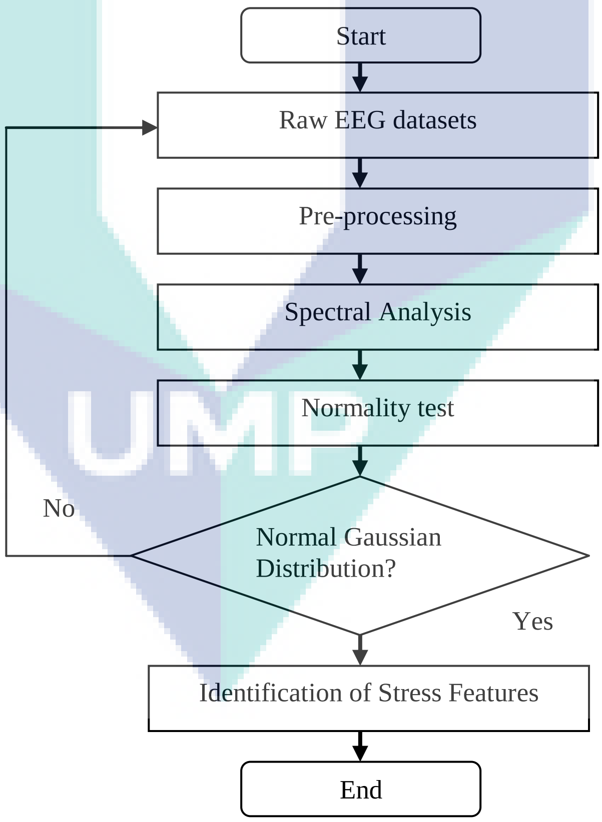


Figure 4.7: Flow Chart for EEG Normality Test

4.6 ANALYSIS OF EEG SIGNALS

This section is the heart of the research as it contains the detailed explanation on how the EEG signals are analyzed and how the features are extracted from the EEG signals. The MATLAB algorithms for analyzing EEG signals can be seen in Appendix F.

4.6.1 Non-parametric Analysis

This research implements non-parametric or spectral analysis of EEG signals to obtain the features that might relate to stress. The analysis are conducted off-line and carried out in frequency-based platform using DFT technique. In computer signal processing packages, the technique is called FFT for the ability of the computer to rapidly calculate and analyze the vast data [7, 60, 112, , 177-178]. The FFT technique produces two spectrums, which are power spectrum and phase spectrum. In addition, this analysis that focuses on the change in the energy of the spectrum due to the change in the cognitive state, and is conducted by analysing the ratio of the ESD and SC for each frequency band and group.

4.6.2 Energy Spectral Density (ESD)

In this research, the ratio of ESD due to the change in cognitive state from doing nothing to answering IQ test questions is used to search for the relative energy for each group. The ESD is determined by dividing the area of the PSD with the frequency range of the bands. First, the amplitude of each band is calculated in terms of Power Spectrum Density after the raw EEG data is converted to spectrum-based signal using FFT. The conversion will produce the PSD of Delta, Theta, Alpha and Beta bands for each group, after which the ESD for each band and group is calculated. In order to determine the RER for each band and group, the total energy for each band is calculated, and the RER of the sub-bands are then calculated by dividing the ESD of the sub-bands with the total energy. Finally, the logarithm is performed to the RER of the sub-bands in order to amplify the value of the RER, especially for the Alpha and Beta bands since the amplitude of the Delta and Theta bands are much higher than that of the Alpha and Beta bands.

4.7 IDENTIFICATION OF STRESS FEATURES

In order to identify the features from EEG signals that might describe the stress level of a person, this study has focused on the following features: (1) asymmetry ratios, (2) natural logarithm of asymmetry ratios and logarithm of RER, (3) entropy of RER, and (4) spectral centroids of RER. Shannon's Entropy (SE) is used to detect any irregularity of the signals and change of the spectral energy, while spectral centroids are used to determine the dominant energy from the groups. The study on asymmetry is performed since the human brain can be divided into the right and left hemispheres. The difference of the brain electrical activities in both hemispheres will create difference cognitive state [7, 12, 21, 26-28] and negative asymmetry can indicate the existence of the stress [22-25, 29, 31]. All the selected features are then used as the input to k-NN classifier for the classification process.

4.7.1 Asymmetry ratio (AR) and Natural log of AR

The asymmetry ratio technique is widely used by researchers to determine the difference cognitive or mental state in certain frequency bands [26-28, 37, 42]. In this study, the score from asymmetry ratio technique is used to reflect the difference of two cognitive states, the condition of not doing anything and answering IQ test questions. The questionnaires for the IQ were set-up based on the standard progressive matrices of IQ questionnaires for visual [73]. The balance cognitive state occurs when the electrical activity in the right hemisphere of the brain equals that of the left hemisphere. If the power spectrum at location Fp1 (left hemisphere) is higher than that at location Fp2 (right hemisphere), brain activity in the right hemisphere is activated, otherwise, brain activity in the left hemisphere is activated. The asymmetry ratio is inversely associated with the activation of brain regions [27-29].

In this study, the asymmetry ratio technique is applied to all frequency bands to determine the asymmetry scores as described by equations 3.11 and 3.12 in Chapter Three. The asymmetry scores are determined from the normalization of asymmetry ratios and natural log asymmetry ratios and is indicated in percentage. The difference score indicates the relative activity of the right and left hemispheres of the human brain, where a negative asymmetry score will indicate the activation of the right hemisphere, while a positive asymmetry score will indicate the activation of the left brain. The negative asymmetry score is associated with negative emotion [1, 3-4, 12, 17-18, 27, 33].

4.7.2 Relative Energy Ratio (RER)

The focus of this study is the difference in energy spectral density due to the change in cognitive state from relax state (not doing anything) to wakefulness state (answering IQ test questionnaires). The RER for each frequency band and group are computed using equations 3.14 to 3.17 as described in Chapter Three. RER is a good feature to indicate the existence of stress due to change in neural activities. The RER values are then fed to k-NN classifier for the classification process.

4.7.3 Spectral Centroids (SC)

SC are widely used in the fields of speech and audio recognition to identify the dominant frequency from noisy speech [133-134]. In this study, SC is used to find the dominant relative energy from the frequency bands and groups, which basically involves the calculation of the spectrogram of the RER at the average frequency of all subjects over the total RER for all subjects as described by equation 3.18 in Chapter Three. Thus, this study uses SC as a class for k-NN classification. First, the SC for each frequency band in one group is calculated, and then the overall SC for all frequency bands is computed. Since four groups are involved in the experiment, there will be four SC which becomes the four classes in the classification process.

4.7.4 Shannon Entropy (SE)

SE is used in this research to detect the pattern that might relate to stress from the four experimental groups, where RER is applied to obtain a value that can be used to indicate a certain pattern of stress. The RER for each frequency band is calculated, after which the mean or average value of SE for each group is determined using equation 3.19 as stated in Chapter Three. The SE is applied to all experimental groups. The classifier performance is calculated based on the three evaluated features, which are asymmetry ratio, natural logarithm of asymmetry ratio and relative energy ratio.

4.8 CLASSIFICATION OF STRESS FEATURES

In k-NN classification processes, the AR and the RER of EEG data are partitioned into training and testing sets, while spectral centroids are used as a target class for the classification process. AR and RER are fed to the classifier's input according to the specified training and testing ratio. In the classification process, other than the input features and the testing and training ratio, the k values are varied in order to obtain the optimal accuracy, sensitivity and specificity. To implement the k-NN classification process in MATLAB, the *knnclassify* function was used to classify the EEG features in term of AR, natural log AR and RER. The performance of the classifier is then validated using cross-validation techniques. This section explains the procedure to classify the selected EEG features starting with classifier configuration, training and testing ratio, and the variation of k value according to the size of the training set. The results of the classification process are discussed in Chapter 5.

4.8.1 k-NN Setting; Distance and Rule

The k-NN classifier is set according to the training and testing ratio, k value, distance and rule. The k-NN classifier distance of Euclidean, City block and Cosine are evaluated, while nearest, random and consensus rule are employed. The default settings of k-NN classifier for distance is Euclidean and that of rule is nearest.

4.8.2 k-NN Setting; Training and Testing Ratio

The classifier is evaluated using two sets of training to testing ratios. For Set 1, the classifier is assessed using 50:50 training to testing ratio, while 70:30 training to testing ratio is used for Set 2. For Set 1, 90 data each are selected for the training and testing groups; however, 126 data for the training group and 54 data for the testing group are selected for Set 2. Different training to testing ratios are employed in the classification process in order to determine the best classification accuracy.

4.8.3 k-NN Setting; k-Value

After selecting the training and testing ratio, the classifier is assessed using different values of k for each set of training to testing ratio. Since the preferred value of k for k-NN classifier is 10, the k value for Set 1 (50:50 ratio) and Set 2 (70:30 ratio) is varied only from 1 to 10. The overall setting for the k-NN classifier are as shown in Table 4.2, Table 4.3 and Table 4.4.

Table 4.2:
k-NN Classification using Asymmetry Ratio as EEG Feature

Features	Training Vs Testing	k value	Distance	Rule
Asymmetry ratio	50:50	1-10	Euclidean, City Block, Cosine	Nearest, Random, Consensus
Asymmetry ratio	70:30	1-10	Euclidean, City Block, Cosine	Nearest, Random, Consensus

Table 4.3:
k-NN Classification using Natural Log Asymmetry Ratio as EEG Feature

Features	Training Vs Testing	k value	Distance	Rule
Natural Log Asymmetry ratio	50:50	1-10	Euclidean, City Block, Cosine	Nearest, Random, Consensus
Natural Log Asymmetry ratio	70:30	1-10	Euclidean, City Block, Cosine	Nearest, Random, Consensus

Table 4.4:
k-NN Classification using Relative Energy Ratio as EEG Feature

Features	Training Vs Testing	k value	Distance	Rule
Relative Energy ratio	50:50	1-10	Euclidean, City Block, Cosine	Nearest, Random, Consensus
Relative Energy ratio	70:30	1-10	Euclidean, City Block, Cosine	Nearest, Random, Consensus

4.9 CLASSIFICATION PERFORMANCE AND VALIDATION

The performance measure of k-NN classification is based on accuracy, sensitivity and specificity, where the performance of each group is determined based on the selected stress features and the spectral centroids of the group. The accuracy, sensitivity and specificity for a 2x2 matrix can be calculated using formulas as described earlier. However, since this research uses EEG datasets that consist of 4x4 matrices, 4x4 confusion matrix is needed in order to measure its classification performance. Table 4.5 shows the table of the confusion matrix, where TP denotes true positives, TN denotes true negatives, FP denotes false positives, FN denote false negatives. Also, “e” denotes a classification error when any feature does not belong to their group and all errors in the table must be calculated. The calculation of accuracy, sensitivity and specificity based on the confusion matrix has been detailed in an earlier chapter. In accordance with the number of stress features, experimental groups, and training to testing ratios, the following six confusion matrices will be produced:

- Confusion Matrix for Asymmetry ratio with Training to Testing datasets ratio of 50:50
- Confusion Matrix for Natural Log Asymmetry ratio with Training to Testing datasets ratio of 50:50
- Confusion Matrix for Relative Energy ratio with Training to Testing datasets ratio of 50:50
- Confusion Matrix for Asymmetry ratio with Training to Testing datasets ratio of 70:30
- Confusion Matrix for Natural Log Asymmetry ratio with Training to Testing datasets ratio of 70:30
- Confusion Matrix for Relative Energy ratio with Training to Testing datasets ratio of 70:30

The classification accuracy, sensitivity and specificity is determined based on the confusion matrices stated above.

Table 4.5:
Confusion Matrix[182-184]

		Predicted	Class		
		A	B	C	D
Actual Class	A	TP_A	eAB	eAC	eAD
(known data)	B	eBA	TP_B	eBC	eBD
	C	eCA	eCB	TP_C	eCD
	D	eDA	eDB	eDC	TP_D

The classification performance is typically assessed by cross-validation technique, which is an established technique to estimate the classifier accuracy and to reduce the bias which deals with the random sampling of training and testing datasets. There are two popular techniques to validate the classifier performance, which are k-fold cross-validation and leave-one-out cross-validation. The size of partition to be used in k-fold validation is dependent on the size the dataset, and it is impractical to use a large partition when the size of the datasets is small. For leave-one-out cross-validation, the first data is selected as the testing data, while the remaining data are used as the training data. Then, the second data is selected as the testing data and the rest of the data become the training data. The process is repeated until the last data in the dataset has been selected as the testing data. These two techniques are applied in this research in order to assess the performance of the classifier.

4.9.1 Classification Error

Besides the classification accuracy, sensitivity and specificity, the classification errors in terms of mean squared error (MSE) are calculated in accordance with the k-NN distance. This technique is implemented to determine which distance produces a low MSE with the highest accuracy.

4.9.2 k-Fold Cross-Validation

This technique creates k-fold partition of the datasets, where k represents the size of the partition of the datasets. Since this study involves datasets with matrix size of 180x4, the 10-fold cross-validation is preferred; the datasets are split into ten partitions, where nine partitions are used for the training set and one partition is for the testing set. The accuracy of the classifier is calculated for every fold, and based on the size of the datasets, 10-fold cross-validation will produce 18 sets of accuracy value. Thus, the overall accuracy is the average of the 18 sets of accuracy values.

4.9.3 Leave-One-Out Cross-Validation

This technique will select one data as the testing data and the rest as the training data. The process begins by selecting data Number 1 as the testing data and ends when data Number 180 is selected as the testing data. Thus, there will be 180 sets of accuracy value. The accuracy is determined from the average value of the 180 sets of accuracy values.

4.10 CLUSTERING OF STRESS FEATURES

Besides the classification of the datasets using k-NN classifier, this research also focuses on fuzzy technology that employs the clustering technique to cluster the stress features from EEG datasets obtained from the four experimental groups. In this study, two type of fuzzy clustering techniques are used, which are FCM and FKM.

4.10.1 Fuzzy C-Means (FCM)

FCM will cluster the feature-vector data into several clusters according to the specified number of cluster. Based on the datasets and number of clusters, FCM analyzes and partitions the datasets into the specified number of clusters, and produces three outputs, namely, center matrix, membership grade matrix and objective function. When the grade is between 0 and 1, this shows that the datasets has partial membership in the cluster. In this study, the datasets is partitioned into four clusters that represents the four different groups involved in the study. The datasets with the specified number of clusters are analysed using the *fcm* function in MATLAB. The formula to implement the function is described by equation 3.34 in Chapter 3. The maximum number of data in the membership matrix obtained from the analysis will serve as an index for partitioning the datasets. The distribution of the datasets in the vicinity of the center matrix can then be observed by plotting the indexed datasets. The classification accuracy is calculated by taking the index of the membership matrix as a class in a k-NN classification.

4.10.2 Fuzzy K-Means (FKM)

Unlike FCM, FKM uses centroids in each cluster, calculated using squared Euclidean distance, to index the location of the data in the cluster, and the index is used to partition the data with the new centroids. The new index obtained from the centroids and clusters is then used as a class in a k-NN classification to determine the accuracy of the classification process. In this study, the FKM is implemented using *kmeans* function in MATLAB. The function is implemented using equation 3.36 in Chapter Three. The input arguments for the function are datasets, the number of clusters and Euclidean distance. This function file will return the index and centroids of the data in the cluster. Based on the index, the new datasets in the clusters is generated, and consequently, the new centroids is produced.

4.11 STATISTICAL ANALYSES: ANOVA and PEARSON CORRELATION STUDY

Statistical analyses, implemented by exporting the features of Social Package for Social Science (SPSS) version 17, are carried out in order to determine the significance of the extracted features within groups and between groups. Besides, it also applied to find the correlation among the selected features. The features consisted of the 180x4 matrices where the rows represent the groups and the columns represent the EEG bands. The analyses are implemented using Analysis of Variance (ANOVA) method and Scatterplot (Pearson correlation) method. In ANOVA, one-way between-groups ANOVA with Post-Hoc tests is selected. It is because the analysis involves the test of one independent variable (EEG features) across the four experimental groups at one time. ANOVA is used to find the significant differences between the groups. In this study, the RER of EEG frequency bands, Asymmetry Ratio of EEG frequency bands and Natural Log Asymmetry Ratio of EEG frequency bands are selected as independent variables while the experimental groups are the dependent variables. Post-Hoc is used to discover where the differences exist or lie. In order to figure out the existence of the relationship between the change of the cognitive state (represented by the EO and EC states) and the PSD of EEG signal for each state, this study uses Pearson correlation. The variables being studied are EEG features in term of RER, AR, natural log of AR with cognitive states represented by the four experimental groups. Therefore, Scatterplot or Pearson correlation study in SPSS is selected. The theoretical explanation of the Pearson correlation study is discussed in Chapter 3.

4.12 STRESS INDEX ASSIGNMENT

The stress index is constructed based on the selected EEG features, spectral centroids and spectral entropy of the features, the ratio of the average energy for each group across the frequency bands and classification results. However, these parameters only apply to the EEG Alpha and Beta power for the change of power in these frequency bands will indicate the change in cerebral state. In this research, the stress index are categorized only into three, which are low stress, moderate stress and high stress. There is no category for stress-free as it is believed that even healthy non-smoking subjects have minimal stress.

4.13 VERIFICATION OF THE STRESS INDICES

Once established, the stress indices are verified using Z-score technique. Z-score calculates the variation of the distribution of the EEG features (asymmetry ratio and energy ratio) in the group. The Z-score value will be produced and will be given a location between -1 and 1, where a score of zero indicates no variation of the individual data in the group. Thus, from the location of the Z-score, the stress indices can be assigned. Besides using Z-score technique, every data in the four experimental groups are verified whether or not they fall into the correct index.

CHAPTER FIVE

RESULTS AND DISCUSSIONS

5.1 INTRODUCTION

The results obtained from the experiments are presented and discussed in this chapter. The study starts with the pre-processing of EEG signals that involves the removal of raw data with unacceptable noise level, removal of artifacts, designing of filter to capture the sub-bands of the EEG signals, performing normality test on the extracted EEG sub-bands, classifying and clustering the EEG features, measuring the classification performance, validating the classifier performance, implementing statistical analysis on the extracted features, producing the stress index, and finally, verifying the stress index.

5.2 EEG SIGNALS PRE-PROCESSING

First, the raw EEG data of each participant was recorded and observed for any abnormality or unacceptable noise. Data that contains unacceptable noise is rejected and replaced with re-generated data. The normal EEG data for the Channel 1 (left channel) and Channel 2 (right channel) is illustrated in Figure 5.1, while the noise pattern is shown by Figure 5.2. The results of applying noise to the original data in order to obtain the re-generated data are shown in Figure 5.3. All the figures are plotted based on the EEG recording for the first and third participants of Group 1. The same procedure is implemented for the data captured from the rest of the subjects. These figures are placed in this chapter to show how the unacceptable and bad data are replaced using data re-generated technique. The raw EEG signals captured from the right and left channels of the first participant, as respectively shown in Figure 5.1 are normal except that certain data on the right channel (Channel 2) and the left channel (Channel 1) exceeded 100 μV as shown by the red markers, which are removed using artifact removal technique. The EEG data were recorded for a duration of three minutes, which is equivalent to 46081 data. It can be seen that the EEG data for Channel 1 and Channel 2 are similar when the subject is in the relax state.

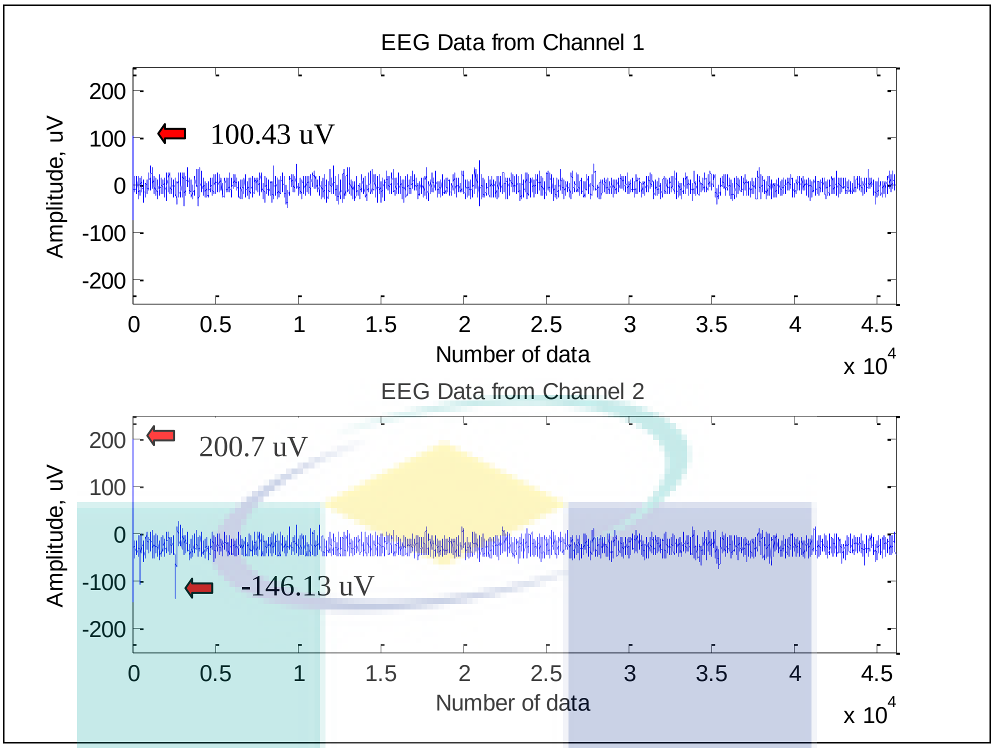


Figure 5.1: Example of EEG Raw Data Recorded from Channel 1 and Channel 2

Figure 5.2 depicts the noise generated using normally distributed pseudorandom numbers technique as described by equation 3.21. The graph consists of the generated noise for the two channels, Channel 1 (Left Channel) and Channel 2 (Right Channel).

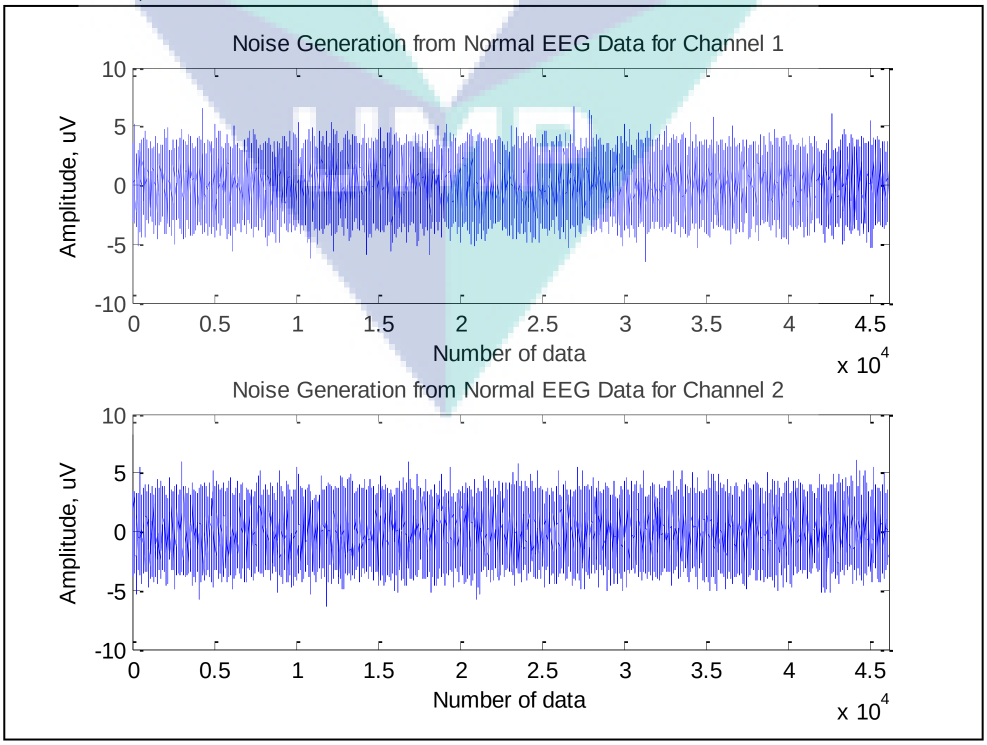


Figure 5.2: Generation of Noise for Both Channels from Normal EEG data

Basically, the generated noise is similar to EEG signals but with very low amplitude, which is less than 10 µV [179]. After applying noise to the original EEG data, the new EEG data is generated as described by equation 3.22 in Chapter 3. According to the formula for re-generating EEG data as described by equation 3.22, the noise is added to the normal or good EEG data with the noise control factor of 1.5. In the process of re-generating EEG data, the maximum and minimum difference between the original and re-generated data must not exceed 10% [179]. Consequently, the number of re-generated data is reduced to 46070 from 46081, and the re-generated data for the left and right channels are shown in Figure 5.3. The length of the 3 minutes recorded EEG data is 46081.

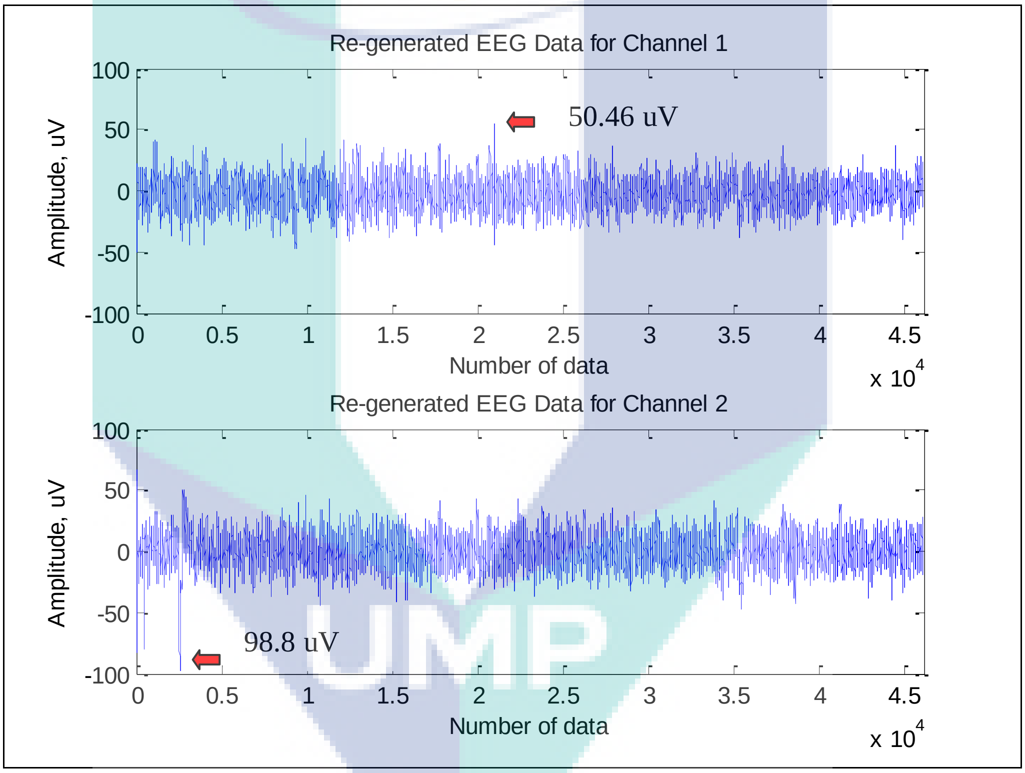


Figure 5.3: Re-generated EEG Data for Channel 1 and Channel 2

Table 5.1 shows the maximum and minimum difference between the original data and re-generated data for Channel 1 and Channel 2. The table shows that only Channel 2 has the minimum difference of 9%. However, the introduction of noise to the original data is smaller, where the maximum noise is approximately 6 µV as shown in Figure 5.2. It is suspected that this is due to the noise (EOG signals) already present in the original data. The maximum and minimum difference between the original and re-generated data is reduced to below 10 % after removing the artifact and filtering the raw EEG signals into its sub-bands.

Table 5.1:
The Percentage Difference between the Original and Re-generated Data

Parameters	Channel 1 (Left)	Channel 2 (Right)
Maximum Difference (%)	33	39
Minimum Difference (%)	22	9

After performing artifact removal and segmenting the EEG signals into their respective frequency bands, the re-generated data are re-checked. The original and re-generated EEG data in term of EEG features such as asymmetry ratio are compared and the results are shown in Table 5.2.

Table 5.2:
The Maximum and Minimum Difference between the Original data and Re-generated Data for both Channels in terms of Asymmetry Ratio

Parameters	Asymmetry ratio (Delta)	Asymmetry ratio (Theta)	Asymmetry ratio (Alpha)	Asymmetry ratio (Beta)
Maximum Difference (%)	0.03	2.80	5.06	8.08
Minimum Difference (%)	1.14	2.41	0.33	4.81

In order for comparison between normal and abnormal distribution of EEG signals, the abnormal EEG data recorded from one of the participants are plotted in Figure 5.4. This type of data was removed and replaced with re-generated data. Figure 5.4 shows that the EEG signals contain a lot of spikes and sometimes have values greater than 100 µV, and this might be due to the eyes movement or blinks (EOG signals). Similar phenomenon is observed in EEG recordings from Channel 2 as shown in Figure 5.4. The maximum EEG values for both Channel 1 and Channel 2 are exceeded 100 µV respectively.

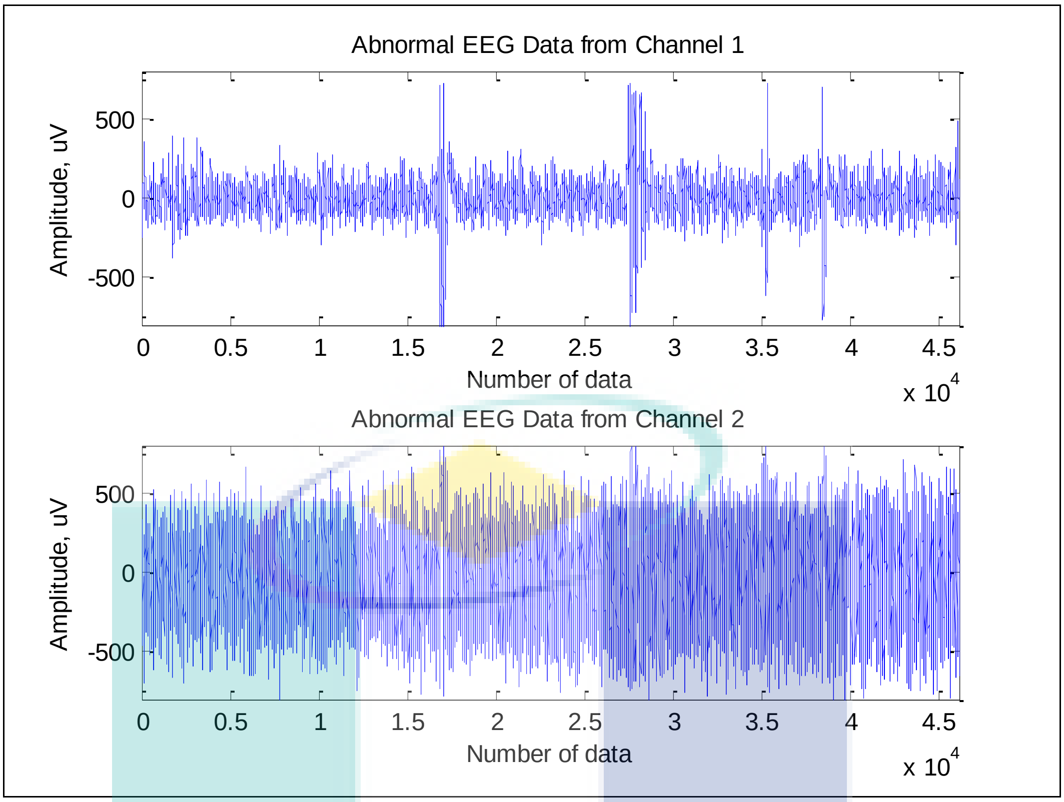


Figure 5.4: The Abnormal EEG Data Recorded from Channel 1 and Channel 2

The type of the recorded EEG data as shown in Figure 5.4 is called corrupted data and therefore, the data need to be removed from the analysis stage. In the experiment stage, there were a total of 37 corrupted EEG data from 180 recorded EEG data which is equivalent to 20.6% and high enough to affect the analysis stage. As shown in Figure 5.5, the magnitude of noise that was added to the original data (abnormal data) to re-generate the data was too small, hence, the addition of noise is not practicable. As illustrated by Figure 5.5, the added noise to the abnormal EEG data is less than ± 10 μV. As the generated noise is of small amplitude, it does not affect the original EEG data, as illustrated by Figure 5.6. According to those figures, the maximum data are greater than 100 μV. Consequently, to replace the corrupted EEG data, the EEG data is re-generated using the normal or good EEG data.

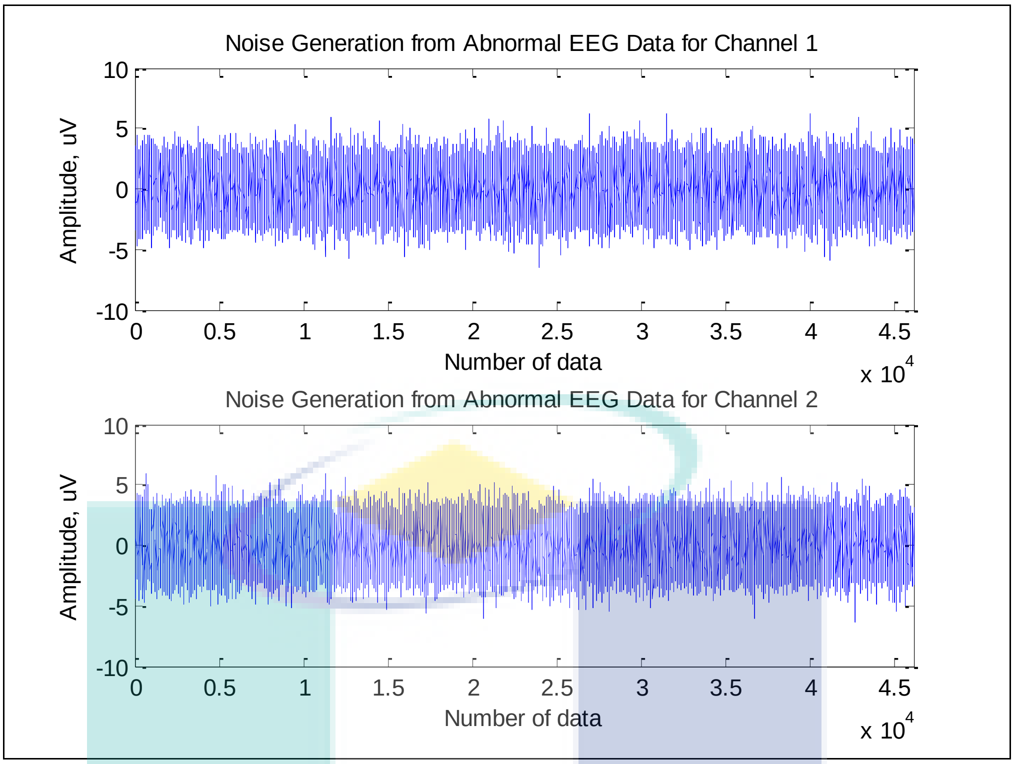


Figure 5.5: Generation of Noise for Both Channels from Abnormal EEG Data

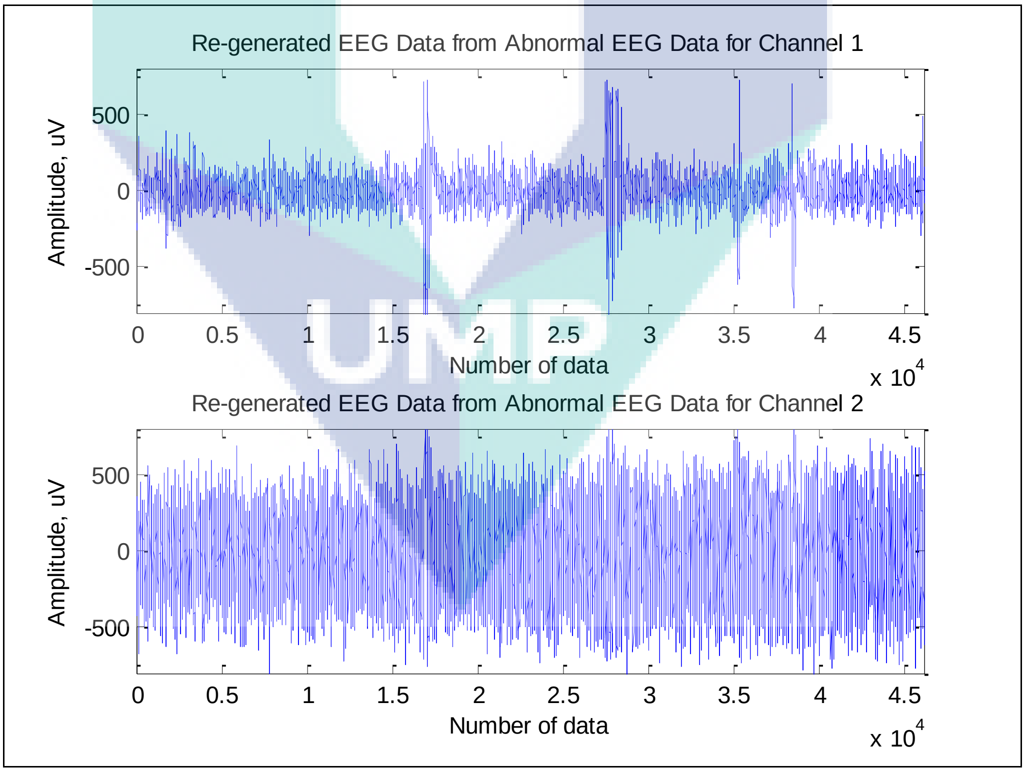


Figure 5.6: The EEG Data Re-generation for Both Channels from Abnormal EEG data

5.2.1 Artifact Removal

This technique is implemented to remove the artifact, especially those generated from eye-movements and eye-blinks. These activities produces ocular artifact or EOG signal, which will reside in the EEG signals since the amplitude and frequency of ocular signals are similar to that of EEG signals. As mentioned earlier, threshold values were assigned in the MATLAB coding to remove EEG data with amplitude above or below 100 µV. The results after applying this threshold value to the EEG data from one of the subjects have been shown in Figure 5.3. It is noted that before undergoing artifact removal, the maximum EEG values for Channel 1 and Channel 2 are 104.43 µV and 200.7 µV, respectively as shown in Figure 1. However, after applying the artifact removal technique, the maximum EEG value for Channel 1 and Channel 2 is reduced to 50.46 µV and 98.3 µV, respectively as shown in Figure 5.3. Consequently, the total number of data are reduced as 12 EEG data with amplitudes above 100 µV have been removed. Similar process is implemented to the rest of EEG data obtained from participants in the experimental groups.

5.2.2 Frequency Filtering

The specifications of the filter employed in this study has been described earlier, and the filtering process is implemented to categorize the EEG signals into their sub-bands. Based on the specifications of the band-pass filter shown in Table 3.1 (Chapter 3) and Figure 4.5 (Chapter 4), the filter meets the cut-off frequency at 0.0039 and 0.2344, as noted in Figure 3.1 (Chapter 3). Besides, the magnitude of passband is lower than -50 dB. The band-pass filter produces the required coefficients to filter the raw EEG signals into Delta, Theta, Alpha and Beta bands, where the frequencies of these bands lie in the range of 0.5 Hz to 30 Hz. A plot of the filter coefficients is illustrated in Figure 3.4 (Chapter 3), where the size of the filter coefficients corresponds to the length of the filter. The characteristics of the designed band-pass filter are shown in Table 5.3.

Table 5.3:
The Characteristics of Window-based Band-pass Filter

Cut-off frequency	Side Lobe Magnitude (dB)	Size of Filter Coefficients	Filter Coefficients
0.0039, 0.2344	< 50 dB	74	-0.09 to 0.13

After the EEG signals have undergone the band-pass filtering process, the right and left channels of the Delta sub-band is illustrated respectively as in Figure 5.7, that of Theta sub-band is as in Figure 5.8, that of Alpha sub-band is as in Figure 5.9, and that of Beta sub-band is as in Figure 5.10. After applying the threshold value, it is noted that all data have amplitude lower than 100 µV. Group 1 consists of subjects that are in a relaxed state and not doing anything while their EEG signals were being recorded, which explains the similarity of the right and left channels of their EEG sub-bands.

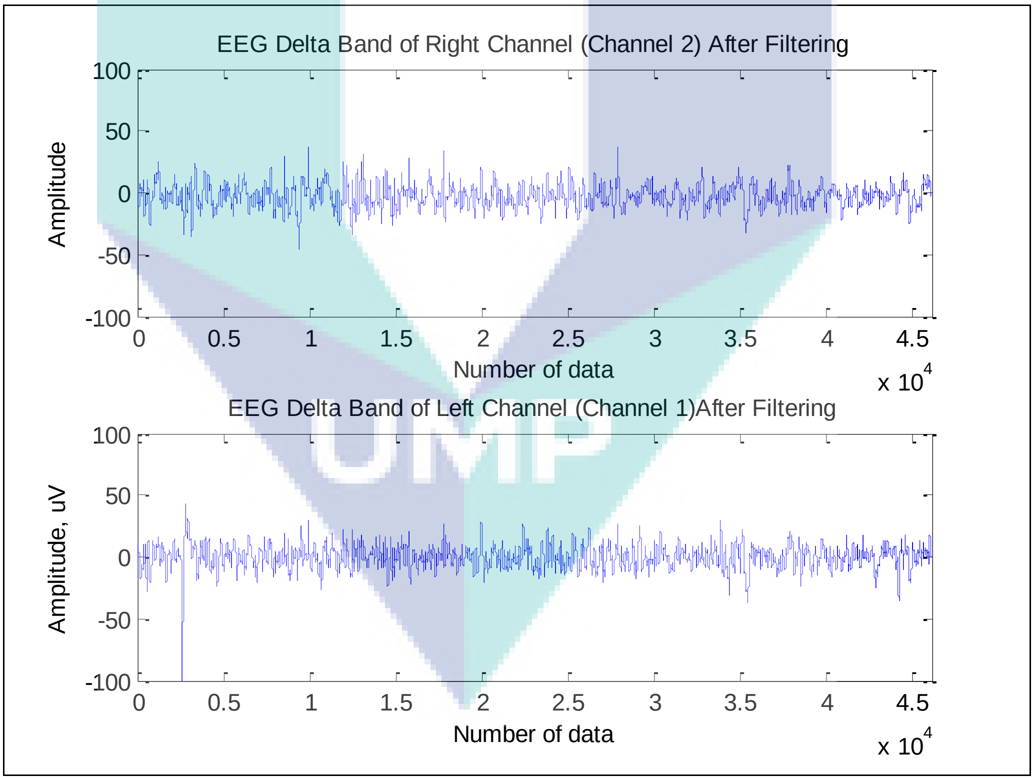


Figure 5.7: EEG Delta band of Right and Left Channel after Filtering

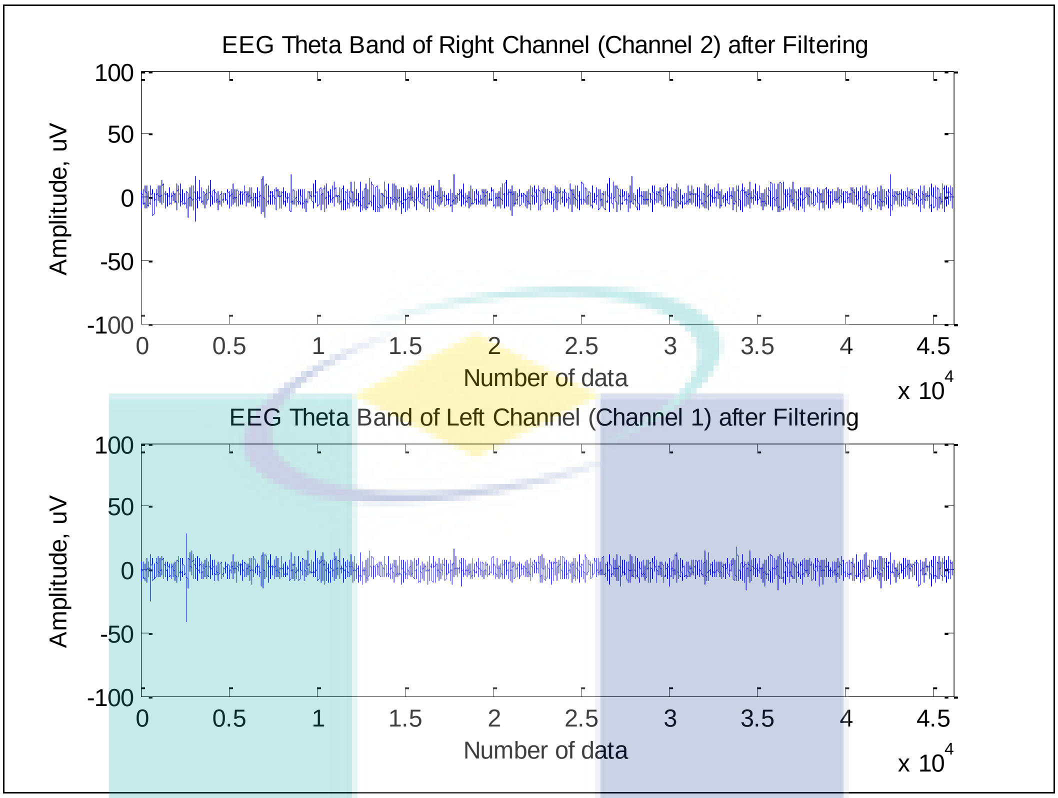


Figure 5.8: EEG Theta band of Right and Left Channel after Filtering

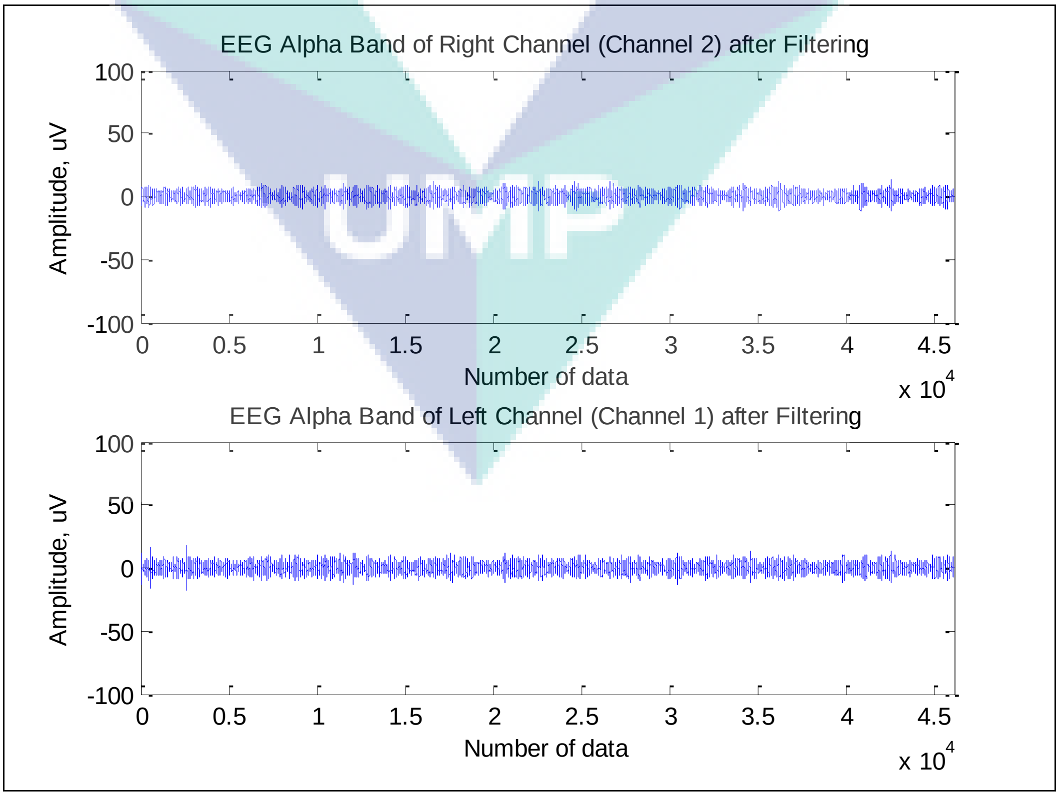


Figure 5.9: EEG Alpha band of Right and Left Channel after Filtering

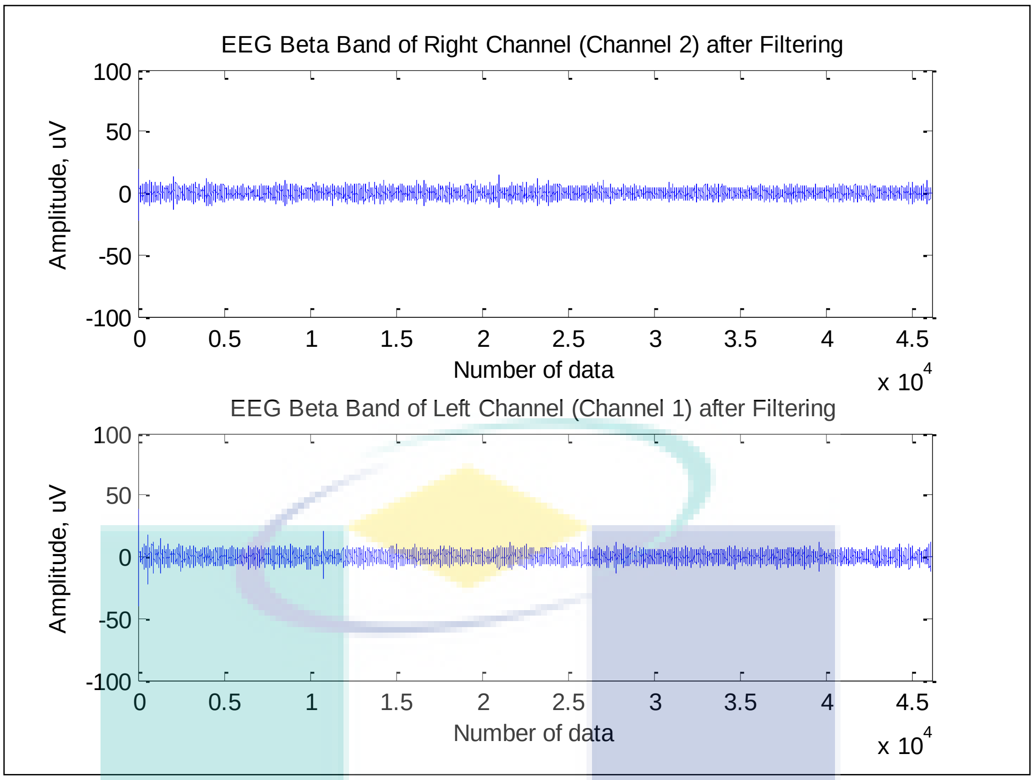


Figure 5.10: EEG Beta band of Right and Left Channel after Filtering

From the figures, in the relaxed cognitive state, it is observed that the EEG signals obtained from the experiments follows the patterns and characteristics of the Delta, Theta, Alpha and Beta sub-bands as described by Table 2.2 in Chapter 2. However, the characteristics changes in the presence of stress as Beta and Theta power increases, while that of Alpha decreases. The most observable parameters are the changes in magnitudes and frequencies for all the bands. It can be seen that the amplitude of the Delta band is higher than that of the Beta band but the frequency of the Delta band is much lower than that of the Beta band, which is the characteristic of EEG frequency bands as-discussed earlier. The characteristic of the EEG data in terms of Power Spectral Density is shown in Table 5.4.

Table 5.4:
EEG Sub-bands in term of Amplitude and Frequency

EEG Sub-bands	Amplitude (μV^2/Hz)	Frequency (Hz)
δ	300 - 1600	0.5 - 4
θ	200 - 600	4 - 8
α	50 - 300	8 - 13
β	10 - 100	13 - 30

5.2.3 Normality Test

Normality test was carried out to determine whether or not the distribution of the recorded data are normal. After performing normality tests to the Power Spectral Density of EEG sub-bands for all experimental groups, histogram plots of ESD ratio for Beta band, Alpha band, Theta band and Delta band are as shown in Figure 5.11 to Figure 5.14 respectively. It is observed that the ESD data for all subjects in the experimental groups have the normal Gaussian distribution, even though the distribution of the data either skewed to the left or skewed to the right. The large skewness is shown by Delta band in Figure 5.22 with long tail to the right. It might caused by the large variation of the EEG power spectrum (ESD ratio) of the Delta band in Channel 1 and Channel 2. The ESD ratio of Beta and Alpha band is slightly skewed to the right (negative skew). Meanwhile, Theta and Delta band are skewed to the left (positive skew).

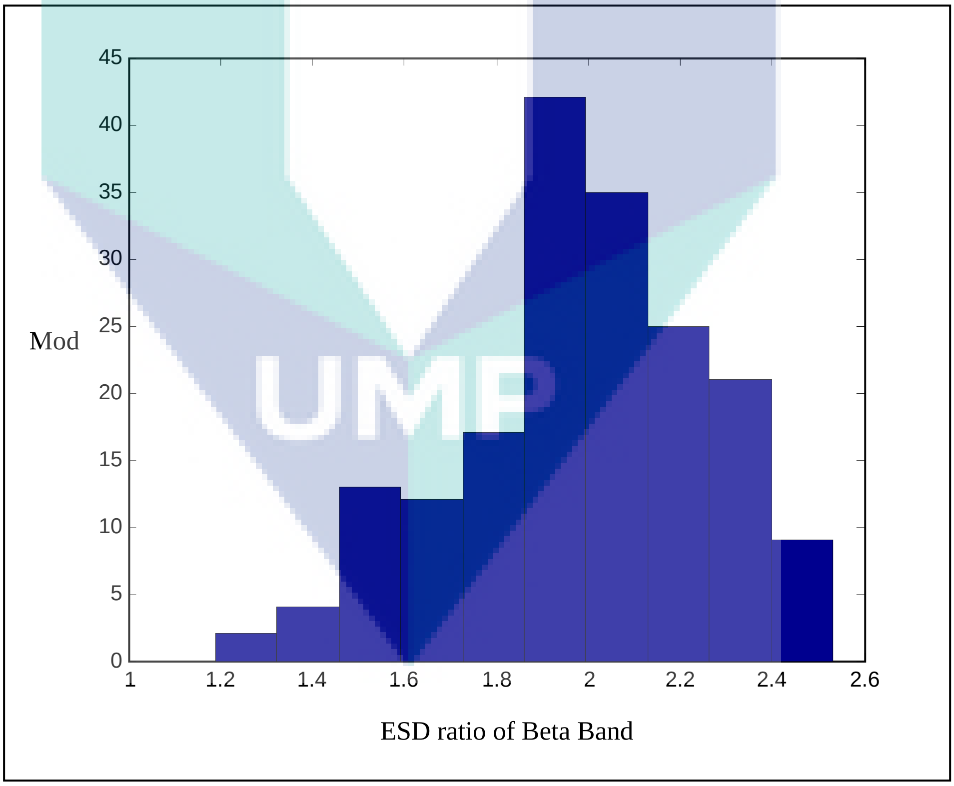


Figure 5.11: Normality Test for ESD ratio of Beta band

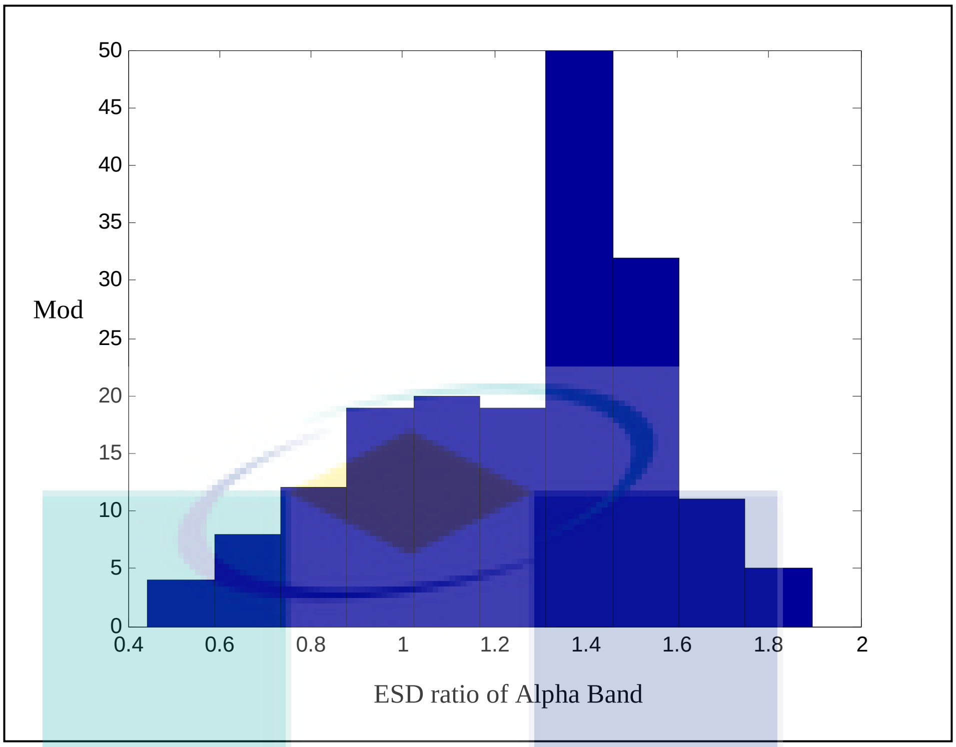


Figure 5.12: Normality Test for ESD ratio of Alpha band

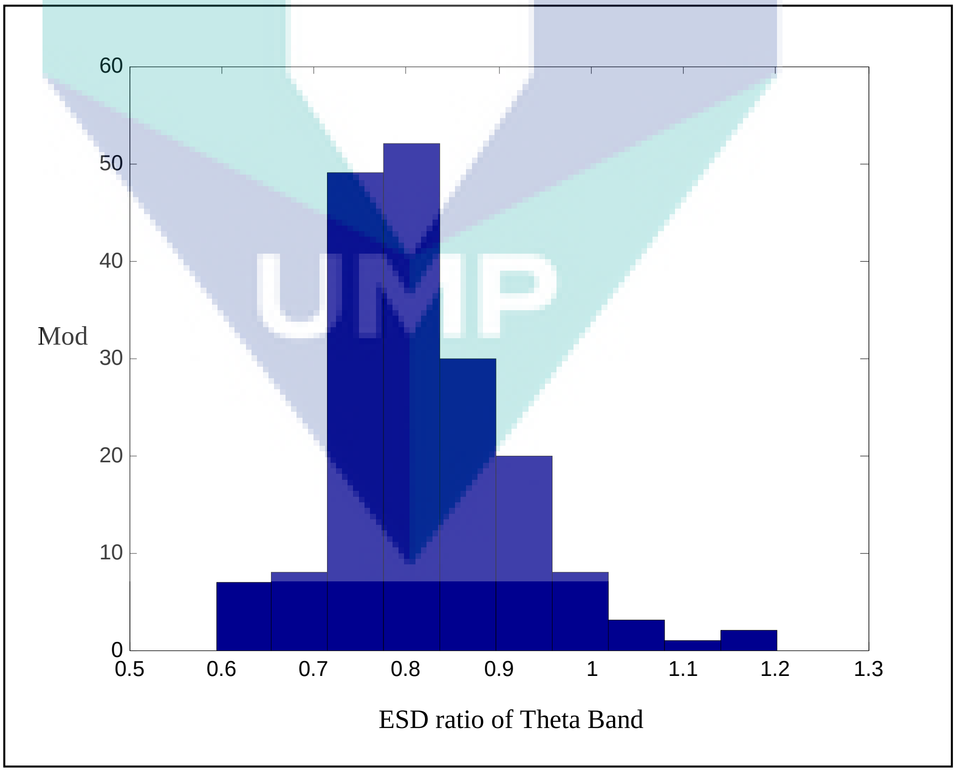


Figure 5.13: Normality Test for ESD ratio of Theta band

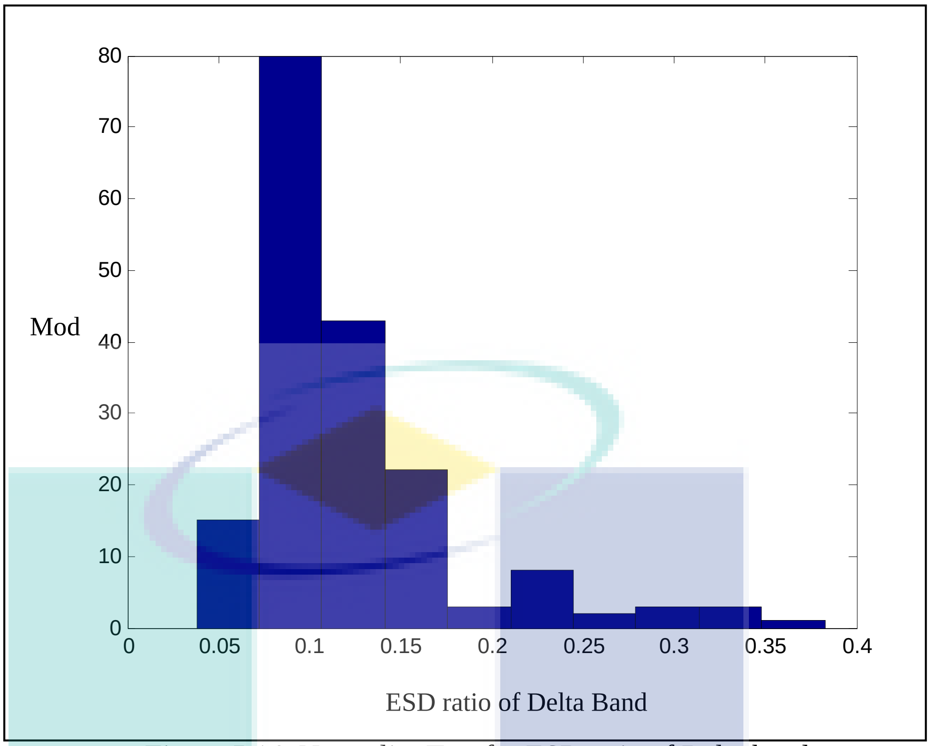


Figure 5.14: Normality Test for ESD ratio of Delta band

5.2.4 Equipment Validation

The EEG device was calibrated according to the voltage, frequency and the offset reading of the electrodes (right channel and left channel). The results of the equipment buy-off and electrodes offsets at different frequencies are as shown in Figure 5.15 to Figure 5.20. The voltages are varied from 100 mV to 500 mV with 50 mV increment. These voltages are generated by Marconi Signal Generator and labeled as VFMSG (Voltage from Marconi Signal Generator). The frequencies of the signal generator are varied from 0.5 Hz to 50 Hz with 1 Hz increment, and a total of nine experiments were conducted for each frequency. The output voltages are tagged as voltage received by g-MOBIlab (VRBG).

Based on the buy-off results shown by Figure 5.23 to Figure 5.28, there are linear relationship between input and output voltages where the voltages received by g-MOBIlab are much higher than the voltages supplied by Marconi Signal Generator except the voltage received by g-MOBIlab at 10 Hz frequency. At those frequency, there is a small different between the input and output voltages even though the voltage received by g-MOBIlab fluctuates across the frequency. The buy-off results also indicate that g-MOBIlab function as voltage amplifier works effectively. In term of frequency, the equipment verification results show that there is no change for the frequencies supplied by Signal Generator and the frequencies received by g-MOBIlab.

However, in term of the average voltage received by EEG electrodes for all experiments, there are an offset of 0.96 as the EEG reading captured by Channel 2 will be 0.96 higher than EEG reading measured by Channel 1, and consequently, the offset is programmed into the MATLAB coding to ensure proper analysis of the captured EEG signals. Regarding the affect of the existence of the measurement offset value, it is not major here. For instance, if the recorded EEG reading from Channel 1 is 100 μV, then, the EEG reading from Channel 2 will be calculated as 104 μV, there is the difference about 4 μV. Another example, 200 μV EEG value recorded by channel 1, will be read about 208 μV by channel 2. As the difference in the EEG reading from both channel will be increased when the amplitude of the EEG signal increase, it will not affect the analysis of the EEG signals in frequency domain, since only clean and normal EEG signals (free of ocular artifact) are used. As discussed in section 4.5 (Pre-processing of EEG Datasets), the threshold value of ± 100 μV was applied in the signal processing algorithm to remove ocular artifact from EEG signals. It means that the signals that are used for spectral analysis are really EEG signals and have amplitude within ± 100 μV.

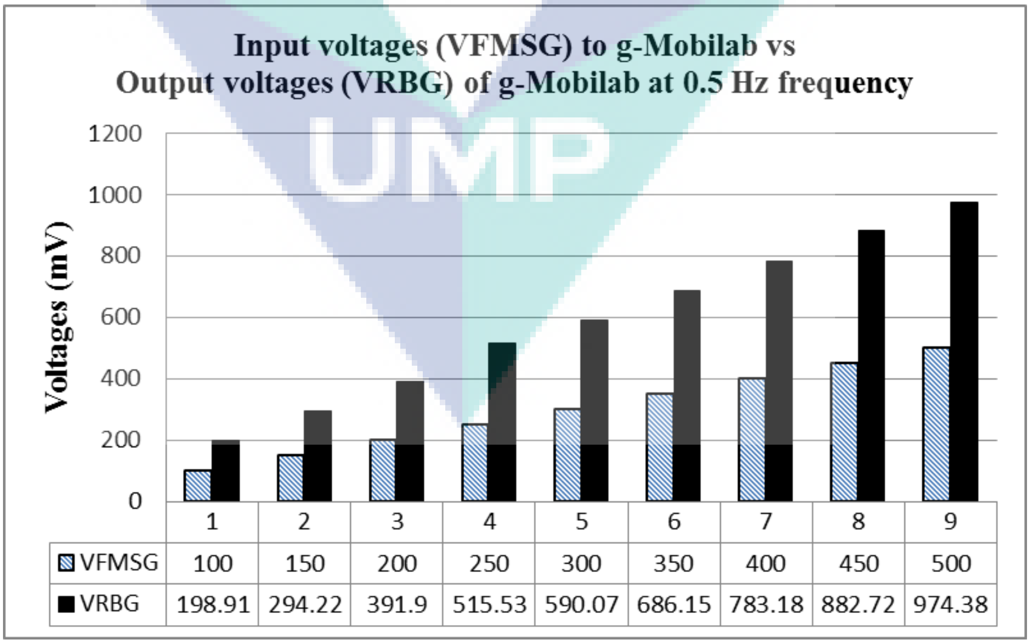


Figure 5.15: Input Voltages versus Output Voltages of EEG Amplifier at 0.5 Hz Frequency

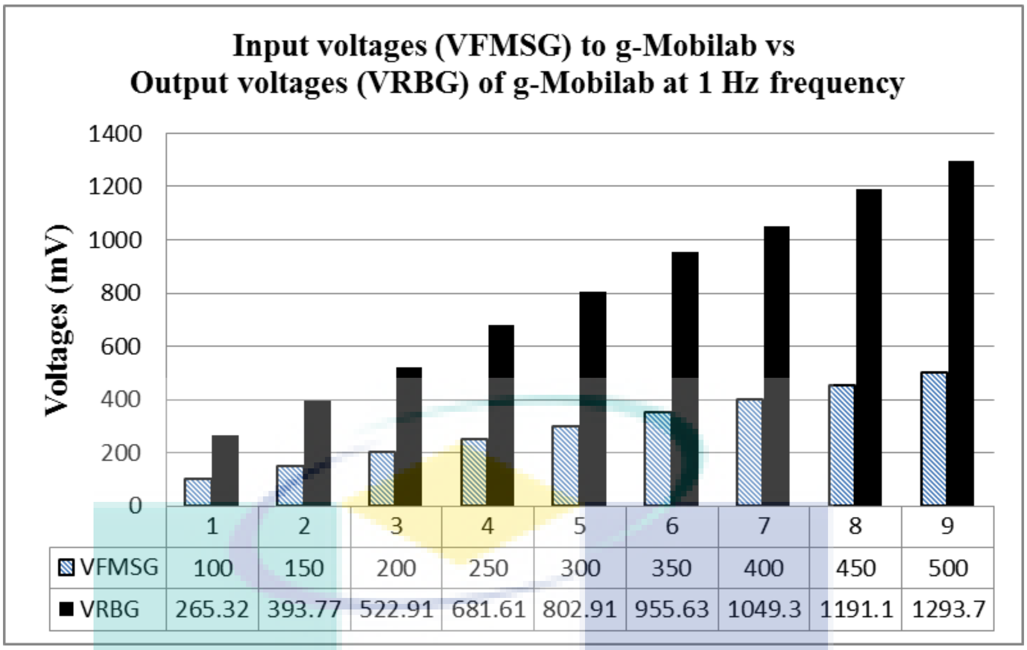


Figure 5.16: Input Voltages versus Output Voltages of EEG Amplifier at 1.0 Hz Frequency

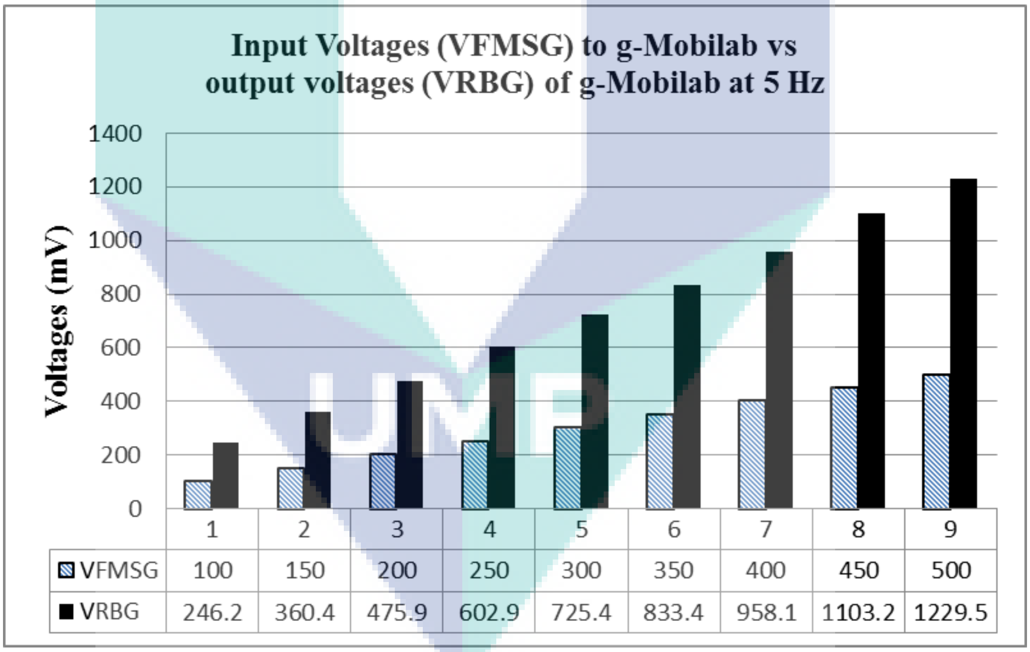


Figure 5.17: Input Voltages versus Output Voltages of EEG Amplifier at 5.0 Hz Frequency

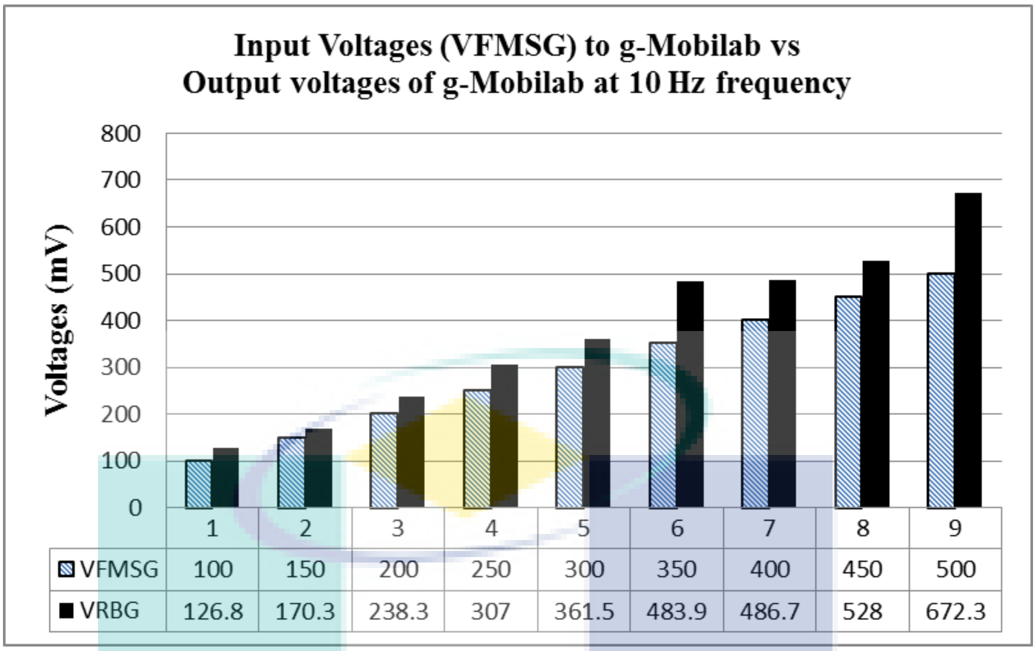


Figure 5.18: Input Voltages versus Output Voltages of EEG Amplifier at 10.0 Hz Frequency

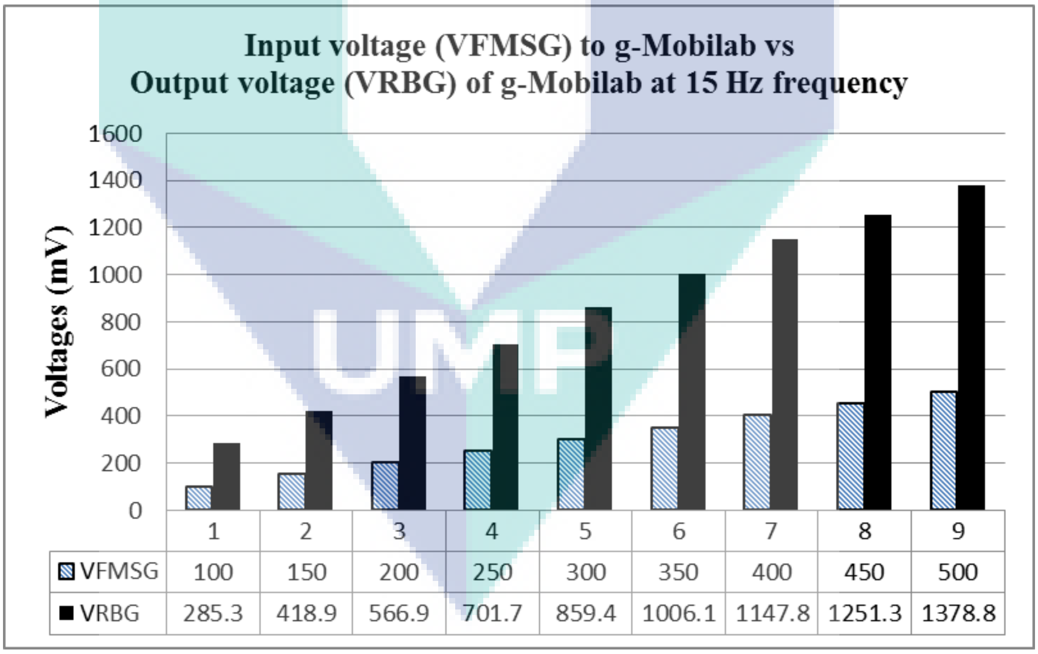


Figure 5.19: Input Voltages versus Output Voltages of EEG Amplifier at 15.0 Hz Frequency

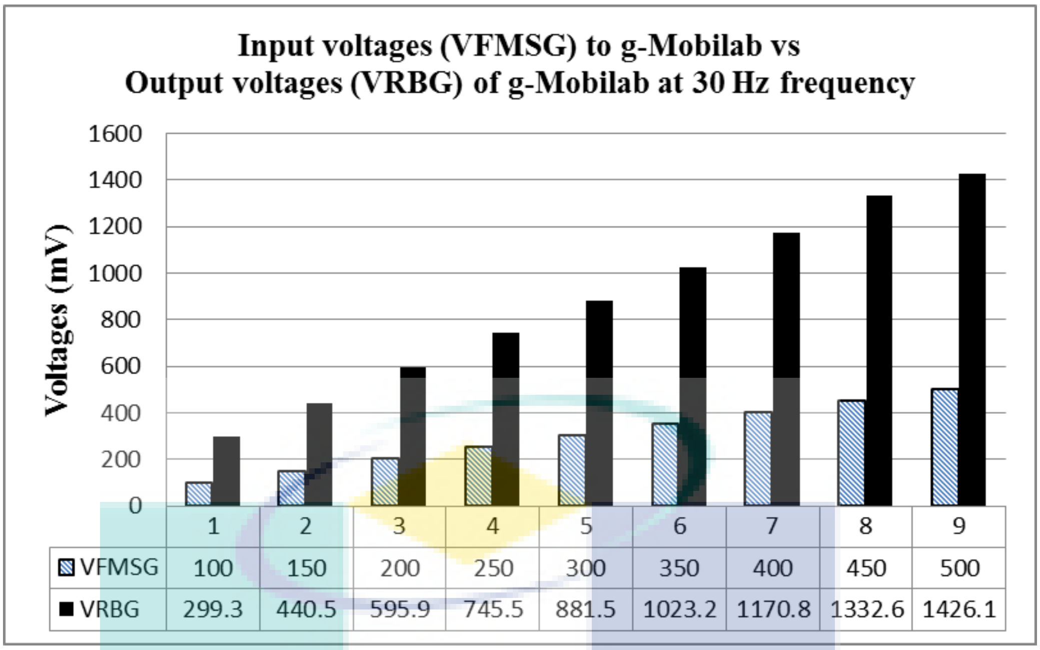


Figure 5.20: Input Voltages versus Output Voltages of EEG Amplifier at 30.0 Hz Frequency

To conclude, g-MOBIlab functionality was verified and it can be used for taking EEG measurement based on the results of the equipment validation. Even though there is an offset in the EEG reading, the offset is small and regardless of voltage fluctuation in certain frequency, there is a linear relationship between the input voltages, VFMSG and the output voltage, VRBG.

5.3 IDENTIFICATION OF STRESS FEATURES

From the spectral analysis of the EEG signals, several parameters were selected as features that might relate to stress. The features chosen are asymmetry ratio, natural log asymmetry ratio, relative energy ratio, spectral centroids and Shannon's entropy. The analysis began by computing ESD using Fourier Transform technique. The block diagram for selecting and classifying the features is illustrated in Figure 5.21. As discussed in previous chapters, spectral centroids act as a target class during the classification process (refer to Appendix C), and entropy acts as an indicator for any abnormal EEG pattern that might occur due to the presence of stress. Meanwhile, asymmetry ratio, log asymmetry ratio and energy ratio were evaluated and tested in order to determine which of the feature can be used to better classify stress. The results of each process shown by the block diagram in Figure 5.21 are discussed in this chapter.

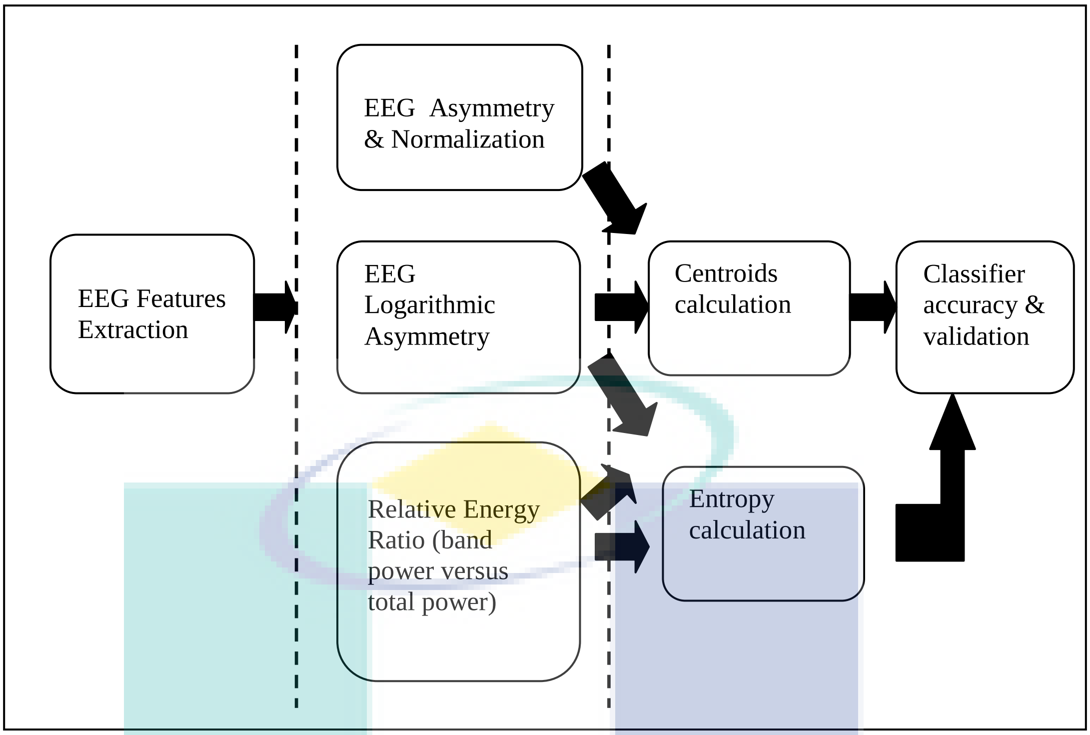


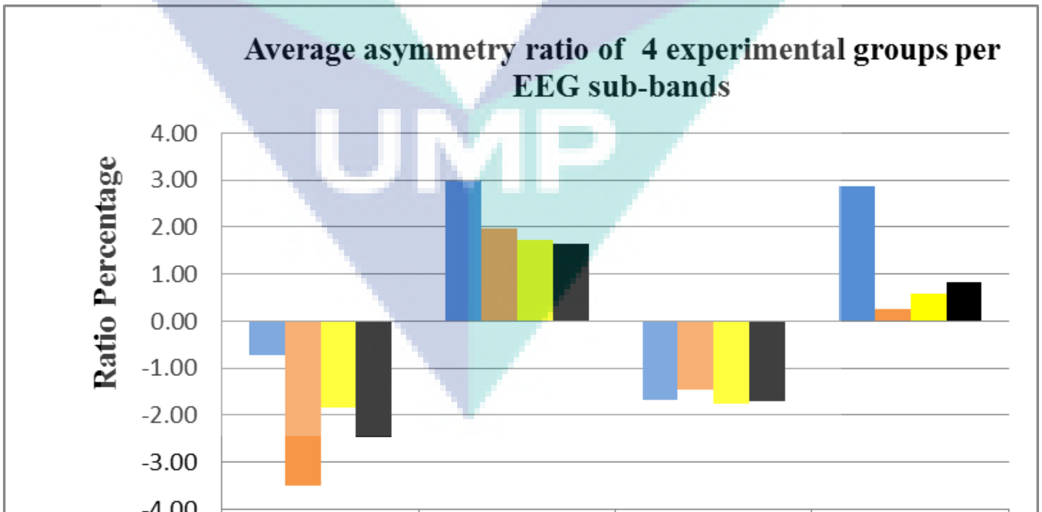
Figure 5.21: Process of Selecting and Classifying EEG Features

5.3.1 Energy Spectral Density (ESD)

After obtaining the ESD values of each EEG frequency sub-band, the ratio of the ESD values is calculated based on the asymmetry and relative energy. The asymmetry ratio is calculated using ln-transformed (natural log) and normalized methods. For the ln-transformed method, the difference between the natural log of the EEG power spectrum at the right hemisphere of the brain and that of the left hemisphere are computed. Meanwhile, for the normalized method, the difference between ESD values at the right hemisphere and left hemisphere are divided by the summation of the ESD values of both hemispheres. For the RER, the relative energy ratio of each frequency band is calculated by dividing the mean value of ESD for each frequency band over the total energy of all frequency bands.

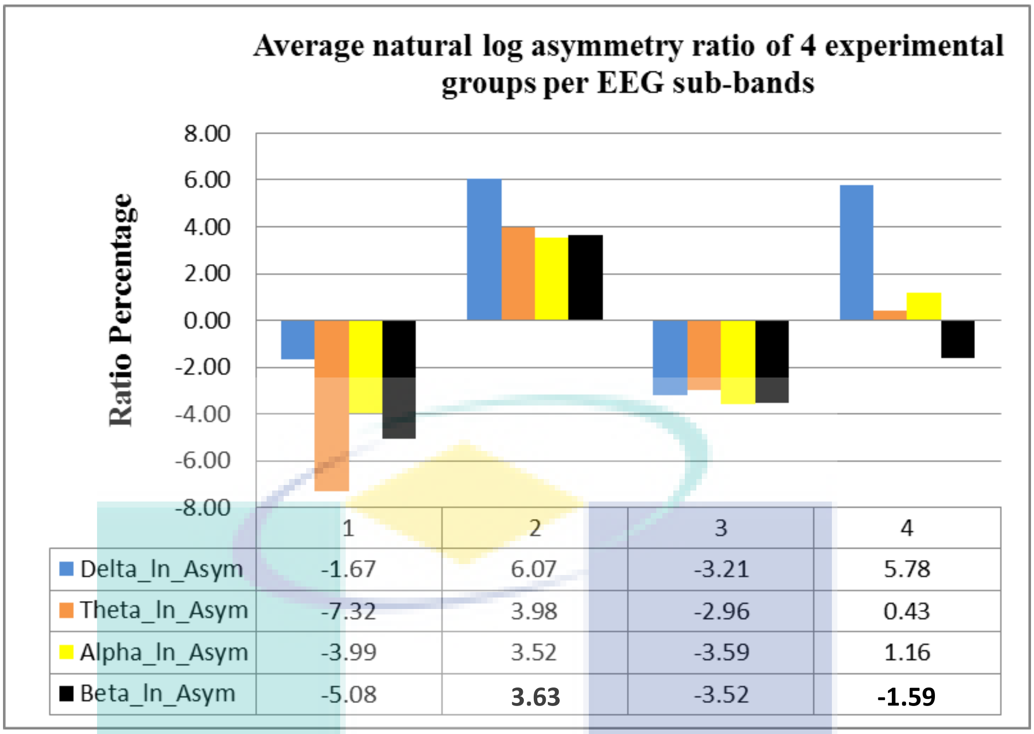
5.3.1.1 Asymmetry Ratio (AR)

Asymmetry ratio is widely used by researchers to look for dominant neural activities, and in this study, an experiment was conducted to calculate the asymmetry ratio of each frequency band of the experimental groups. The asymmetry ratio obtained using ln-transformed and normalized methods for each participant are shown in Appendix C. Figure 5.22 and Figure 5.23 illustrate the mean value of asymmetry ratios and natural logarithm of asymmetry ratios in terms of percentage of each experimental group. Almost the same pattern of average asymmetry distribution shown by both figures except for asymmetry ratio of Beta band in Group 4. The average asymmetry ratios in Group 1 are slightly lower than Group 2. It might due to the different asymmetry formula to calculate asymmetry ratio as explained in Chapter 3 (Section 3.3.3). From both figures, it is observed that for all EEG frequency sub-bands, Group 1 that consists of participants in a relaxed state have negative asymmetry ratios, while that of Group 2 who have to perform IQ tests have positive asymmetry ratios. In addition, the asymmetry ratio of Group 3 (before rotational bed intervention) is similar to that of Group 1, while the asymmetry ratio of Group 4 (after rotational bed intervention) is similar to that of Group 2 except for the Beta band. These results tally with that of previous studies on asymmetry ratio where the brain cognitive state in the right hemisphere will be activated when a person had to perform some tasks. The neural activities in the right hemisphere of the human brain are more dominant than that of the left hemisphere. Hence, from the results obtained, asymmetry ratio can be considered as one of the EEG features.



	1	2	3	4
Delta_Asym	-0.73	2.96	-1.68	2.88
Theta_Asym	-3.50	1.96	-1.45	0.25
Alpha_Asym	-1.85	1.72	-1.77	0.59
Beta_Asym	-2.48	**1.64**	-1.70	**0.84**

Figure 5.22: Asymmetry Ratio of ESD for all Experimental Groups per EEG sub-bands



	1	2	3	4
Delta_ln_Asym	-1.67	6.07	-3.21	5.78
Theta_ln_Asym	-7.32	3.98	-2.96	0.43
Alpha_ln_Asym	-3.99	3.52	-3.59	1.16
Beta_ln_Asym	-5.08	**3.63**	-3.52	**-1.59**

Figure 5.23: Natural Log Asymmetry ratio of ESD for all Experimental Groups per EEG sub-bands

5.3.1.2 Relative Energy Ratio (RER)

Before proceeding to compute RER, the Fourier transform of the EEG signals captured from 140 subjects and the computation of the power spectrum from 180 EEG datasets are examined. Examples of the formation of the power spectrum for Delta band, Theta band, Alpha band and Beta band are illustrated in Figure 5.24 to Figure 5.27 respectively. The process of spectral analysis of EEG signals began by determining the FFT value from the filtered EEG signals, and then follow with the calculation of the PSD. In order to observe the distribution of the power of EEG sub-bands across the frequency, the power spectral density estimation of the Welch spectrum was also calculated. The graphs were plotted from the analysis of the recorded EEG signals from one of the participants in Group 1. In all the figures, the plots in the first row represent the Fourier transform of the EEG signals for the right channel (first column) and left channel (second column). In the second row are the results of squaring the values of the Fourier transform in the first row, which results in the doubling of its amplitude. The third row represents the power spectral density of the EEG signals after applying the length of the FFT (Fast Fourier Transform), window and window overlapping to the values found in the second row. The forth row of the figures describes the estimation of power spectral density using Welch spectrum. By using this spectrum, the power spectral density of the EEG signals will be calculated based on the Decibel and frequency. It is used to verify the calculation of the PSD using FFT. The MATLAB *psd* and *Welch* spectrum functions are employed to produce the plots.

Figure 5.24 illustrates how the plot of the Delta band from right and left channel, which is located approximately in the range of 0 Hz to 4 Hz and peaking around 2 Hz, is obtained based on the filter specifications, band specifications and type of window used in the filter specifications. The peak amplitude of the PSD for this band is approximately 3000 μV^2 / Hz, which tallies with the characteristic of the Delta band as described by Table 2.2 in Chapter 2 and Figure 3.6 in Chapter 3. As decribed by Table 2.2 and Figure 3.6, Delta band has the characteristics of higher amplitude at the low frequency compared to the other EEG frequency bands. It is observed that the amplitude of the PSD of the left channel is a little bit higher than right channel. However, the plot of the Welch spectrum (the last row of the plot) shows about the same distribution of the Delta power for both channels.

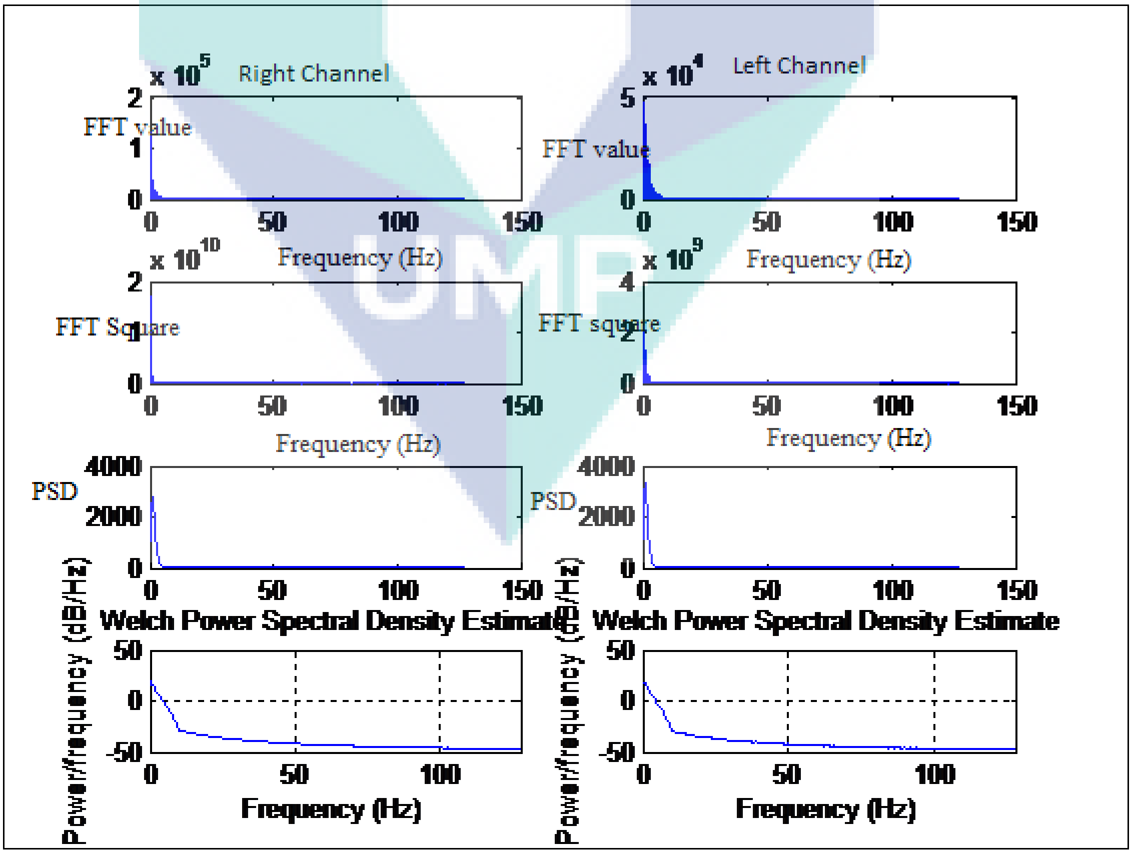


Figure 5.24: The Spectral Analysis to Produce Delta band

Figure 5.25 illustrates how the plot of the Theta band is obtained, which is located approximately in the range of 0 Hz to 8 Hz and peaking around 6 Hz. The peak amplitude of the power spectral density for this band is approximately 300 – 400 μV^2 / Hz, which tallies with the characteristic of the Theta band. As decribed by Table 2.2 and Figure 3.6 in Chapter 2 and Chapter 3 respectively, Theta band has the characteristics of high amplitude at the low frequency. But, the amplitude of Theta band is lower than amplitude of Delta band. It is observed that the there is wide distribution of the FFT plot for the left channel compared to right channel. It indicates the high cognitive activities at the left hemisphere of the human brain. The FFT plots show that there are also high variation of Theta power at left channel compared to right channel. However, the PSD value and Welch power spectrum for both channels are similar.

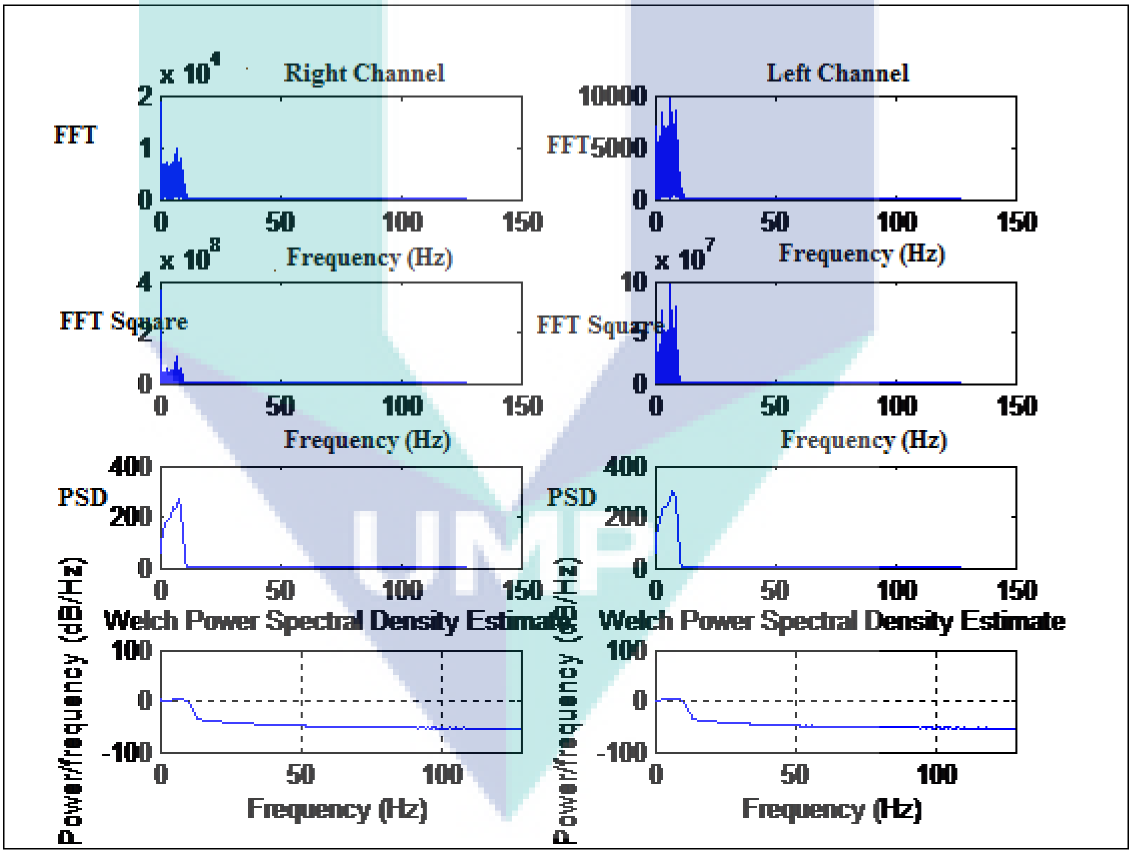


Figure 5.25: The Spectral Analysis to Produce Theta band

Figure 5.26 illustrates how the plot of the Alpha band is obtained, which is located approximately in the range of 8 Hz to 13 Hz and peaking around 10 Hz. The peak amplitude of the power spectral density for this band is approximately 500 μV^2 / Hz. As decribed by Table 2.2 and Figure 3.6 in Chapter 2 and Chapter 3 respectively, Alpha band has the characteristics of low amplitude at the high frequency. From Figure 5.26, both channels show about the same amplitude and distribution of FFT and PSD value which indicate the balanced brain electrical activities at both brain hemispheres. The balanced electrical activities at both channels are confirmed by the plot of Welch spectrum. These plots of Alpha band also tallies with the characteristic of the electrical activities of Alpha band which indicating that the person with this pattern of Alpha band might experience a relax situation.

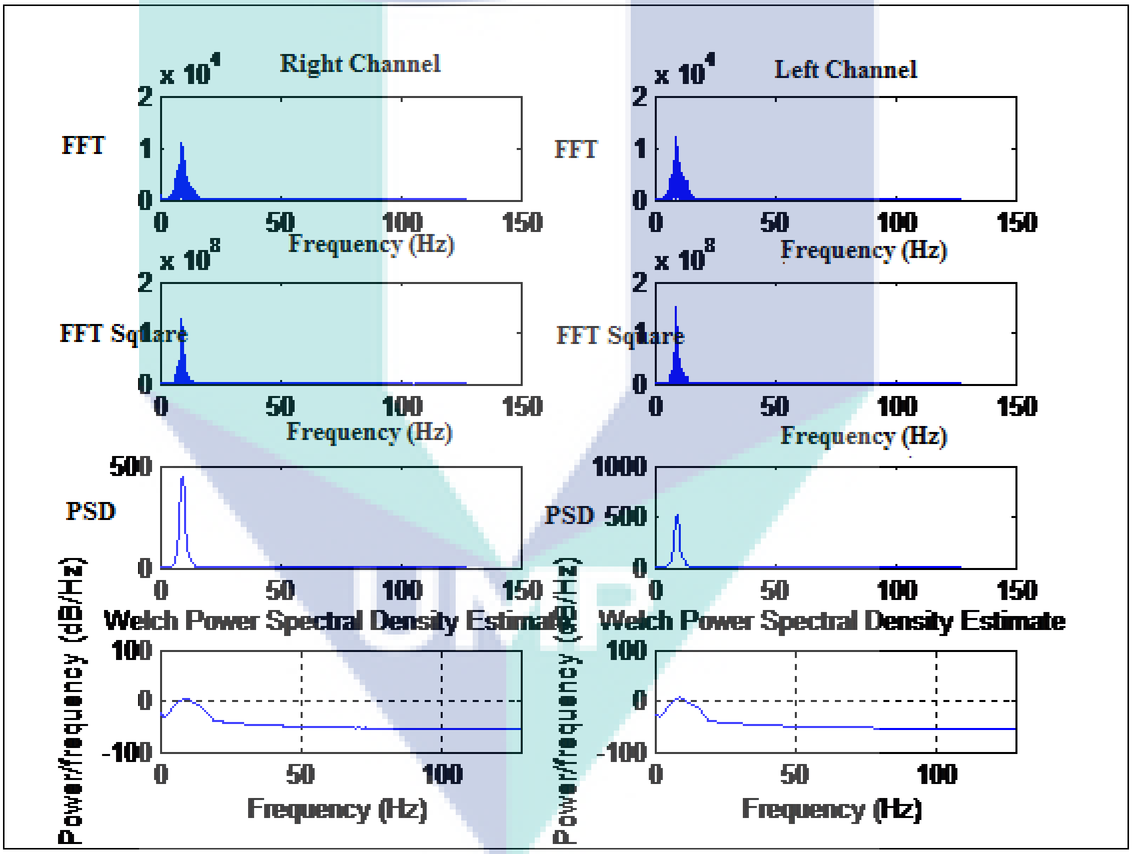


Figure 5.26: The Spectral Analysis to Produce Alpha band

Figure 5.27 illustrates how the plot of the Beta band is obtained, which is located approximately in the range of 13 Hz to 30 Hz and peaking around 21 Hz. The peak amplitude of the power spectral density for this band is approximately 100 μV^2 / Hz. As decribed by Table 2.2 and Figure 3.6 in Chapter 2 and Chapter 3 respectively, Beta band has the characteristics of lower amplitude at the high frequency compared to the other EEG frequency bands. As illustrated by Figure 5.27, there are wide distribution for FFT and PSD plots of Beta band. Beside confirming the distribution of power spectrum and energy of Beta band versus its range of frequency, the noisy FFT plots for both channel indicate that there are high variation of the EEG Beta band at both channels. The distribution of the PSD plots of the EEG for the left channel is wider than right channel even though the peak value of the PSD plots are almost the same for both channels. It might also indicates that the energy of the EEG Beta band at the left hemisphere of human brain is higher than the right hemisphere of human brain. Thus, there are high electrical activities at that area which might indicate the left frontal activation of human brain has occurred. According to the literature review, the person with this pattern of EEG Beta band might experience positive emotion such as happy or excited [23-28]. The plot of the Welch spectrum for both channels are also similar with amplitude around -50dB/Hz which confirming the used of *Hamming* window in designing filter to produce the required filter coefficients to extract the EEG frequency bands from their raw data.

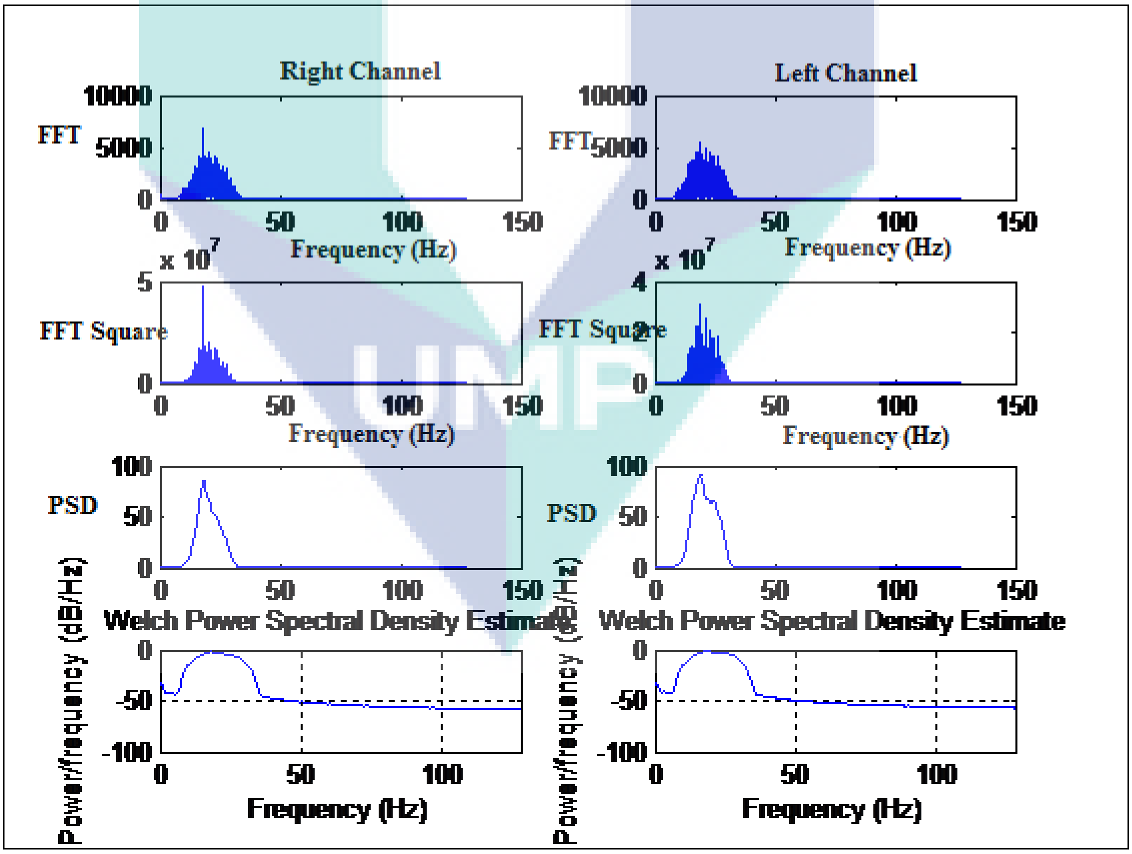


Figure 5.27: The Spectral Analysis to Produce Beta band

In order to observe the pattern of the RER per group as the selected EEG features, the RER per group versus total energy of the group for all frequency bands is plotted as illustrated by Figure 5.28. The presence of stress is indicated by the changes in the power of Alpha band and Beta band, especially that of Group 1 and Group 2, while the power of Delta band and Theta band is not taken into account as all data are captured when the subjects are fully awake. In Group 1, where the participants were in a relaxed state and not doing anything, the RER of Alpha band is higher than that of the Beta band. On the contrary, in Group 2, where the participants have to perform the IQ test, the RER of Beta band is higher than that of the Alpha band. However, the RER of Alpha band of Group 3 and Group 4 is much higher than that of Group 2, while the RER of the Beta band of Group 3 and Group 4 are slightly lower than than that of Group 2. This indicates that the participants in Group 3 and Group 4 are in a more relaxed state than those in Group 2. The trend of RER of Alpha band and Beta band of Group 2 might suggest the change in cerebral activities due to the existence of stressors or stress factors. On contrary, the high trend of RER of Alpha band in Group 4 might indicate subjects have experienced relax condition after undergone intervention (rotational bed) in which the value of RER almost similar with the value of RER of Alpha band in Group 1. In addition, Figure 5.28 indicates that the variation of energy in EEG frequency bands exists among the 4 experimental groups.

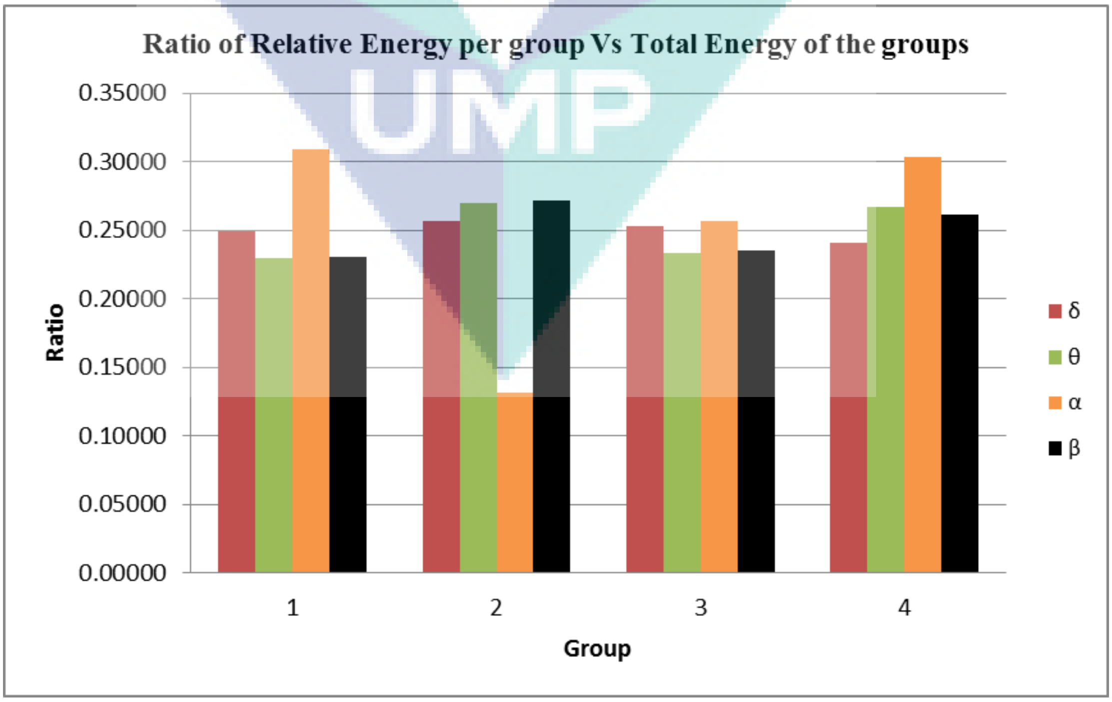


Figure 5.28: RER per group versus Total Energy of the Group

5.3.2 Spectral Centroids

Spectral Centroids (SC) is one of the selected EEG features that can be used to search the pattern of stress from EEG signals. The value of SC then will be used as classification target in classification process. As discussed earlier, SC are applied on the relative energy ratio of all EEG frequency bands to identify the dominant energy among the groups when they change their task from doing nothing to perform IQ test , and the results of the application are shown in Figure 5.29. Again, as the main focus of this study is to examine for changes in Beta band and Alpha band, it is observed that the centroids value of the Beta band in Group 2 is higher than that of the Alpha band of the same group. It is because the cognitive activities of the subjects in Group 2 increase when performing IQ test. Meanwhile, the centroid values of the Beta band in Group 1, Group 3 and Group 4 are lower than that of the Alpha band of the same group. The centroid value of Alpha band in Group 1 is higher compared to the other groups. It indicates that the subjects in Group 1 is really in relax condition. The results of the experiment indicate that application of spectral centroids on EEG power spectrum can be used to identify human stress.

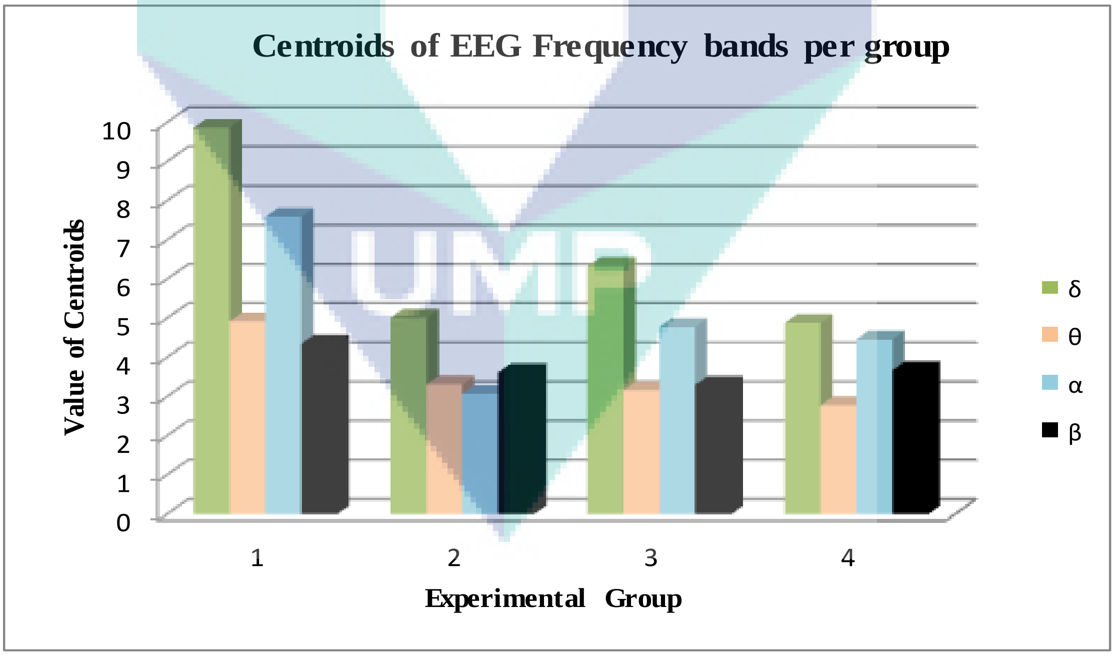


Figure 5.29: The Spectral Centroids of the Experimental Groups

Meanwhile, the average values of the spectral centroids per group across the EEG frequency bands are tabulated in Table 5.5 and its histograms are plotted as shown in Figure 5.30. It is observed that there is a significant trend of average centroids per group as illustrated by Figure 5.30. It is noted that the average centroids value of Group 2 is the lowest among the groups, and can be considered as one of the significant trend even though the difference is only 0.3 (20%) from that of Group 1. The highest centroids come from Group 1 (**9.87**) and the lowest centroids come from Group 4 (**2.78**), while the average centroids value for Group 1, Group 3 and Group 4 is approximately the same with the difference among them ranging from 0.04 (3%) to 0.07 (7%). The low value of average centroids of Group 2 represents the EO or alert cognitive state of the subjects in those group where they are required to perform IQ test. Meanwhile, the similar pattern of average centroids of the other 3 groups indicate that the cognitive state of the subjects on those groups are in EC state or relax state. Hence, the analysis of the experimental results confirm that SC can be employed to indicate the status of the cognitive state of a person effectively.

Table 5.5:
The Value of Centroids Across EEG Frequency bands for All Groups

Group	Centroids of δ	Centroids of θ	Centroids of α	Centroids of β	Average Centroids
1	**9.87**	4.92	7.59	4.35	1.53
2	5.02	3.31	**3.07**	**3.62**	**1.23**
3	6.33	3.17	4.78	3.30	1.49
4	4.88	**2.78**	4.46	3.67	1.46

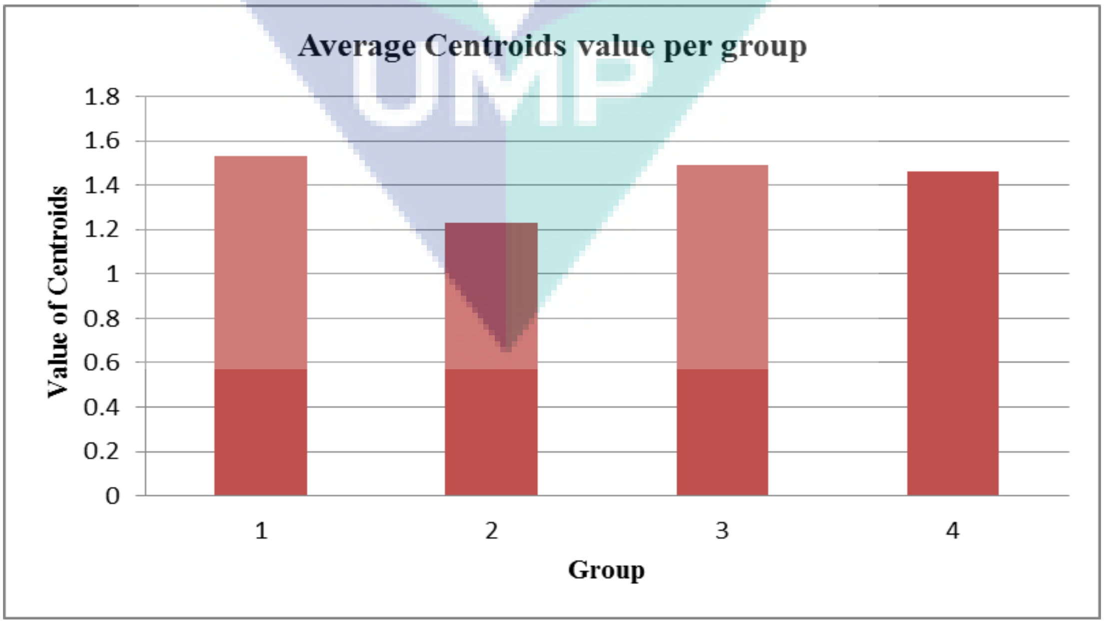


Figure 5.30: The Histogram plot of the Average Centroids for All Groups

5.3.3 Shannon Entropy

The results of applying SE to the RER across the frequency bands of every group are shown in Figure 5.31. It is observed that using SE has produced results that are approximately the same as that obtained when using SC, where Group 2 has the lowest entropy value (**0.29**) compared to other groups. Meanwhile, Group 4 has the highest entropy value (**0.33**). Entropy is applied to the RER of each group to detect any abnormal pattern in the EEG signals, and the low spectral entropy value of Group 2 indicates irregularity in the signals. The analysis result indicates that, besides the average centroid value of the groups, SE can be considered as a reliable feature to detect the change in the distribution of energy of EEG signals which might suggest the presence of stress. Therefore, both SC and SE can be used to detect the changes of the cognitive state from relax to alert state or vice versa.

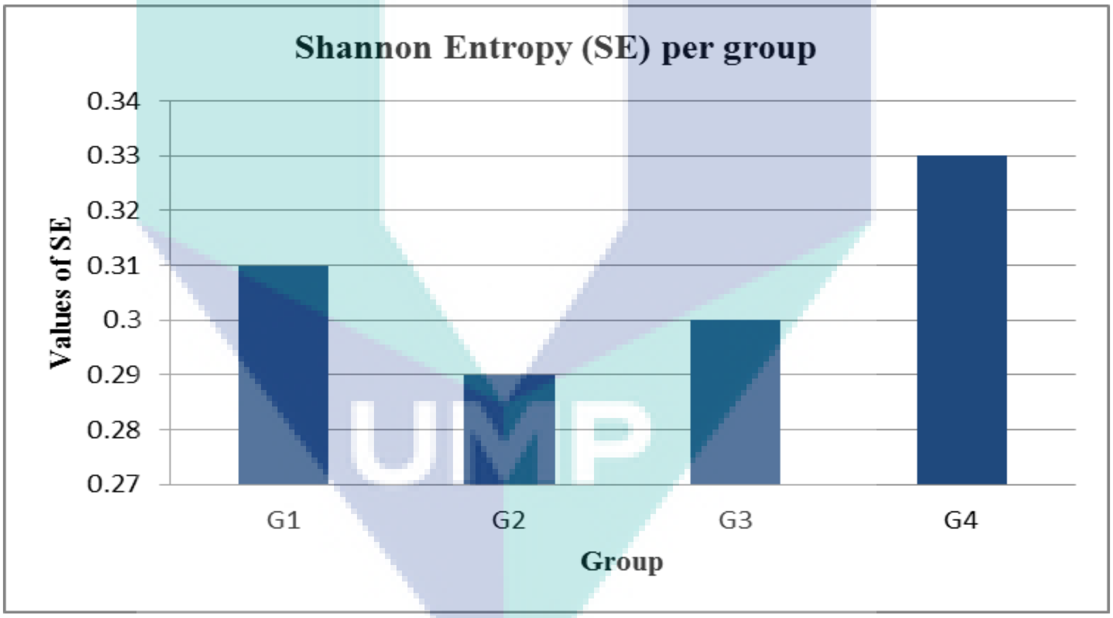


Figure 5.31: The Histogram plot of Shannon Entropy for All Groups

5.4 RESULTS OF THE k-NN CLASSIFICATION OF THE STRESS FEATURES

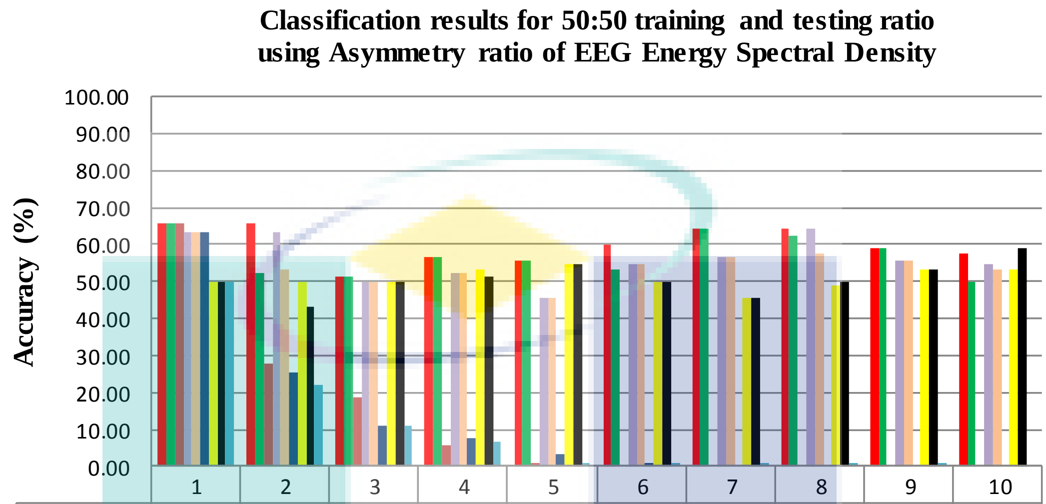
The selected stress features are classified using k-NN classifier to determine which of the experimental groups might be affected by stress. In order to obtain the optimal classification results, the k-NN classifier need to be configured according to the value of k, testing to training datasets ratio, and distance and rule. As described before, three types of features are fed to k-NN classifier, which are normalization of asymmetry of EEG energy spectral density, natural log of asymmetry of EEG energy spectral density and relative energy ratio of EEG signals. The k-NN configurations are then tuned according to ratio of training to testing datasets and the setting of distance and rule, while varying k value in each experiment. With 180 data available from the four experimental groups, the k value is varied from 1 to 90 using 50:50 training to testing ratio, while the value is varied from 1 to 126 for the 70:30 training to testing ratio. However, as explained in earlier chapters, the k value of the k-NN classification is typically small, hence only the k value from 1 to 10 are shown in this thesis. The overall evaluation results obtained from varying the k value are as attached in Appendix C. The discussion in this section starts with the experimental results of k-NN classification using normalization of asymmetry of ESD with the testing and training ratio of 50:50, and is followed with the configuration of the k-NN setting according to the distance and rule.

5.4.1 Classification using Normalization of Asymmetry and Natural Log of Asymmetry of ESD

This section discusses the results of the k-NN classification using different EEG features and the affect on the classifier performance when the training to testing ratios is varied. Besides, the affect of changing the classifier configurations towards classification results are also discussed.

5.4.1.1 50:50 Training to Testing Ratio

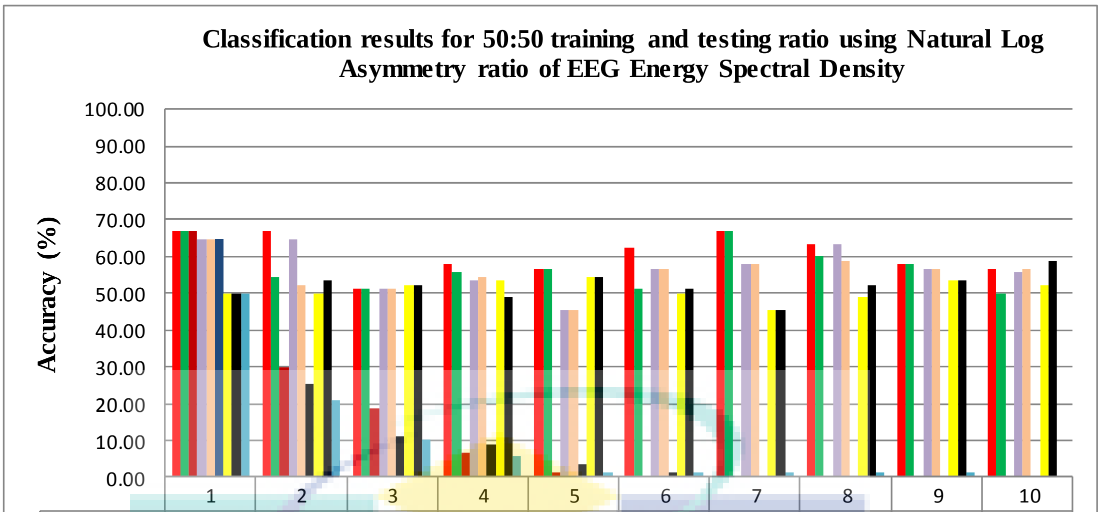
The accuracy of the stress features classification using k-NN classifier with normalization of asymmetry of ESD as an input feature for 50:50 training to testing ratio are shown in Figure 5.32, where the maximum classification accuracy obtained is 65.56% at k = 1 and k = 2. Even though asymmetry can be used to detect the dominant brain activity in the right hemisphere or left hemisphere of the human brain, the classification accuracy using this feature as input is unsatisfactory. Since the testing to training ratio also play a vital role in determining the classifier accuracy, the experiment is conducted at different training to testing ratios. For the next experiment, the classification is implemented at the same training to testing ratio but using natural log of asymmetry of ESD as the features to be classified, and the results obtained are shown in Figure 5.33.



	1	2	3	4	5	6	7	8	9	10
Euc&Near	65.56	65.56	51.11	56.67	55.56	60.00	64.44	64.44	58.89	57.78
Euc&Rand	65.56	52.22	51.11	56.67	55.56	53.33	64.44	62.22	58.89	50.00
Euc&Con	65.56	27.78	18.89	5.56	1.11	0.00	0.00	0.00	0.00	0.00
City&Near	63.33	63.33	50.00	52.22	45.56	54.44	56.67	64.44	55.56	54.44
City&Rand	63.33	53.33	50.00	52.22	45.56	54.44	56.67	57.78	55.56	53.33
City&Con	63.33	25.56	11.11	7.78	3.33	1.11	0.00	0.00	0.00	0.00
Cos&Near	50.00	50.00	50.00	53.33	54.44	50.00	45.56	48.89	53.33	53.33
Cos&Rand	50.00	43.33	50.00	51.11	54.44	50.00	45.56	50.00	53.33	58.89
Cos&Con	50.00	22.22	11.11	6.67	1.11	1.11	1.11	1.11	1.11	0.00

Figure 5.32: The Classification Results using Asymmetry of ESD at 50:50 Training and Testing ratio

The use of natural log of asymmetry of ESD has slightly improved the classification results to 66.67% accuracy at k=1, k=2 and k=7 as shown in Figure 5.33. Again, the highest classification results were obtained at the classification distance and rule of Euclidean and nearest respectively. However, the classification accuracy was still low even though the classification accuracy has improved by 1%. Therefore, the features are evaluated at 70:30 training to testing ratio in order to improve the classification accuracy.

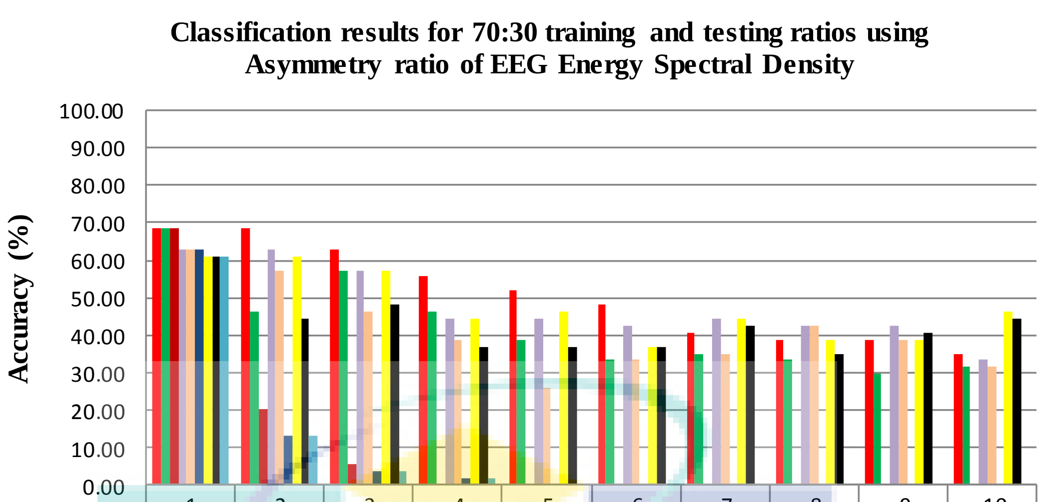


	1	2	3	4	5	6	7	8	9	10
Euc&Near	66.67	66.67	51.11	57.78	56.67	62.22	66.67	63.33	57.78	56.67
Euc&Rand	66.67	54.44	51.11	55.56	56.67	51.11	66.67	60.00	57.78	50.00
Euc&Con	66.67	30.00	18.89	6.67	1.11	0.00	0.00	0.00	0.00	0.00
City&Near	64.44	64.44	51.11	53.33	45.56	56.67	57.78	63.33	56.67	55.56
City&Rand	64.44	52.22	51.11	54.44	45.56	56.67	57.78	58.89	56.67	56.67
City&Con	64.44	25.56	11.11	8.89	3.33	1.11	0.00	0.00	0.00	0.00
Cos&Near	50.00	50.00	52.22	53.33	54.44	50.00	45.56	48.89	53.33	52.22
Cos&Rand	50.00	53.33	52.22	48.89	54.44	51.11	45.56	52.22	53.33	58.89
Cos&Con	50.00	21.11	10.00	5.56	1.11	1.11	1.11	1.11	1.11	0.00

Figure 5.33: The Classification Results using Natural Log Asymmetry of ESD at 50:50 Training and Testing ratio

5.4.1.2 70:30 Training to Testing Ratio

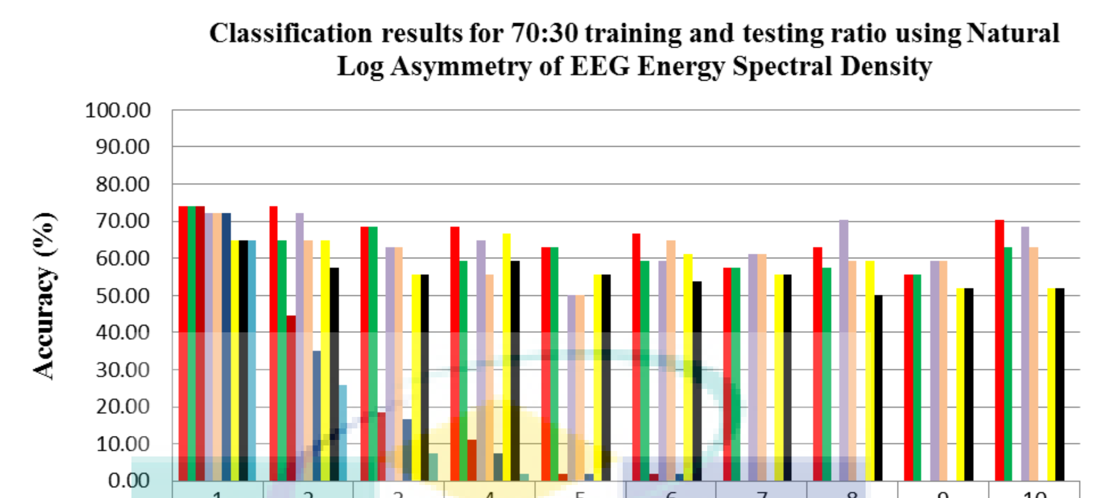
Figure 5.34 depicts the increase of classification accuracy by approximately 2% from 66.67% to 68.52% when 70:30 training to testing ratio is used while still employing normalization of asymmetry of EEG's ESD as the input to the classifier. The results indicate that, besides the features, the large percentage of training datasets has assisted the increase in the classifier performance. However, the highest classification was obtained at classifier's distance of Cosine and rule of nearest, random and consensus. It is normal to achieve the highest classification results at a small value of k, which is a characteristic of k-NN classifier.



	1	2	3	4	5	6	7	8	9	10
Euc&Near	68.52	68.52	62.96	55.56	51.85	48.15	40.74	38.89	38.89	35.19
Euc&Rand	68.52	46.30	57.41	46.30	38.89	33.33	35.19	33.33	29.63	31.48
Euc&Con	68.52	20.37	5.56	0.00	0.00	0.00	0.00	0.00	0.00	0.00
City&Near	62.96	62.96	57.41	44.44	44.44	42.59	44.44	42.59	42.59	33.33
City&Rand	62.96	57.41	46.30	38.89	25.93	33.33	35.19	42.59	38.89	31.48
City&Con	62.96	12.96	3.70	1.85	0.00	0.00	0.00	0.00	0.00	0.00
Cos&Near	61.11	61.11	57.41	44.44	46.30	37.04	44.44	38.89	38.89	46.30
Cos&Rand	61.11	44.44	48.15	37.04	37.04	37.04	42.59	35.19	40.74	44.44
Cos&Con	61.11	12.96	3.70	1.85	0.00	0.00	0.00	0.00	0.00	0.00

Figure 5.34: The Classification Results using Asymmetry of ESD at 70:30 Training and Testing ratio

However, with the same training and testing ratio of 70:30, the classification using natural log of asymmetry of ESD outperforms the classification using normalization of asymmetry of ESD as the input to the classifier. The classification results indicate the improvement of the accuracy of 5%. In this case, the classification accuracy has increased to 74% as shown in Figure 5.35, and again, the highest classification results were obtained at the classification distance of Euclidean and rule of nearest, random and consensus. This indicates that the selection of the feature and selection of the training and test ratio are very crucial in improving the classification accuracy.



	1	2	3	4	5	6	7	8	9	10
Euc&Near	74.07	74.07	68.52	68.52	62.96	66.67	57.41	62.96	55.56	70.37
Euc&Rand	74.07	64.81	68.52	59.26	62.96	59.26	57.41	57.41	55.56	62.96
Euc&Con	74.07	44.44	18.52	11.11	1.85	1.85	0.00	0.00	0.00	0.00
City&Near	72.22	72.22	62.96	64.81	50.00	59.26	61.11	70.37	59.26	68.52
City&Rand	72.22	64.81	62.96	55.56	50.00	64.81	61.11	59.26	59.26	62.96
City&Con	72.22	35.19	16.67	7.41	1.85	1.85	0.00	0.00	0.00	0.00
Cos&Near	64.81	64.81	55.56	66.67	55.56	61.11	55.56	59.26	51.85	51.85
Cos&Rand	64.81	57.41	55.56	59.26	55.56	53.70	55.56	50.00	51.85	51.85
Cos&Con	64.81	25.93	7.41	1.85	0.00	0.00	0.00	0.00	0.00	0.00

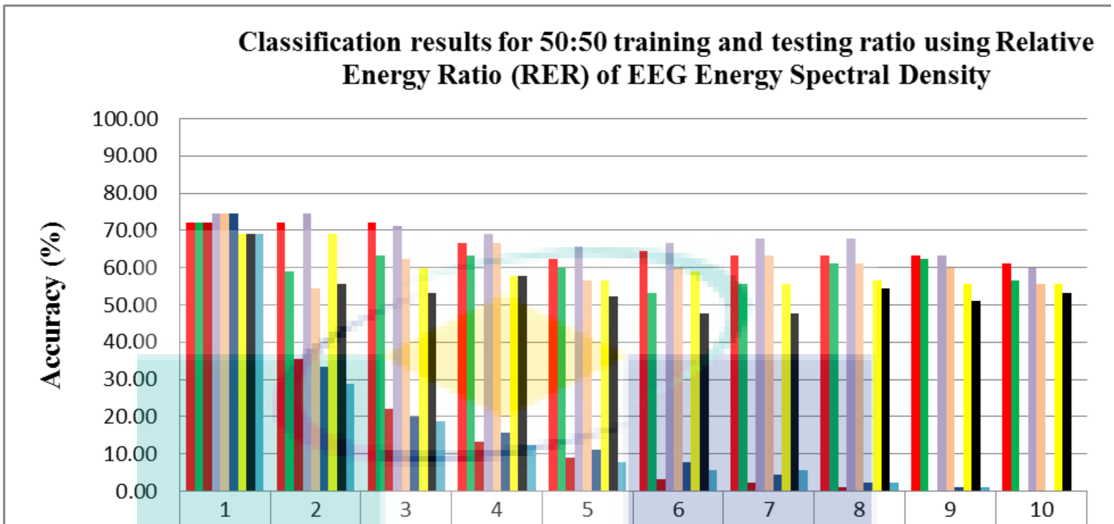
Figure 5.35: The Classification Results using Natural Log Asymmetry of ESD at 70:30 Training and Testing ratio

5.4.2 Classification using Relative Energy Ratio

Besides asymmetry of ESD, relative energy ratio of ESD is fed to the classifier in order to investigate how this feature can affect the classifier performance. The same procedure were applied where the data is split into 50:50 and 70:30 of training to testing ratio, while the classifier distance and rule are varied.

5.4.2.1 50:50 Training to Testing Ratio

Figure 5.36 shows the results when relative energy ratio with 50:50 training to testing ratio are applied to the k-NN classifier. A classification accuracy of 74.44 % was achieved by configuring the classifier with distance of Cityblock distance, and rule of nearest, random and consensus where k remains low. It is a slim improvement compared to the classification results using Natural Log Asymmery at 70:30 testing and training ratios. When the classifier's distance and rule are respectively changed to Euclidean and nearest, the classification performance dropped to 72%.

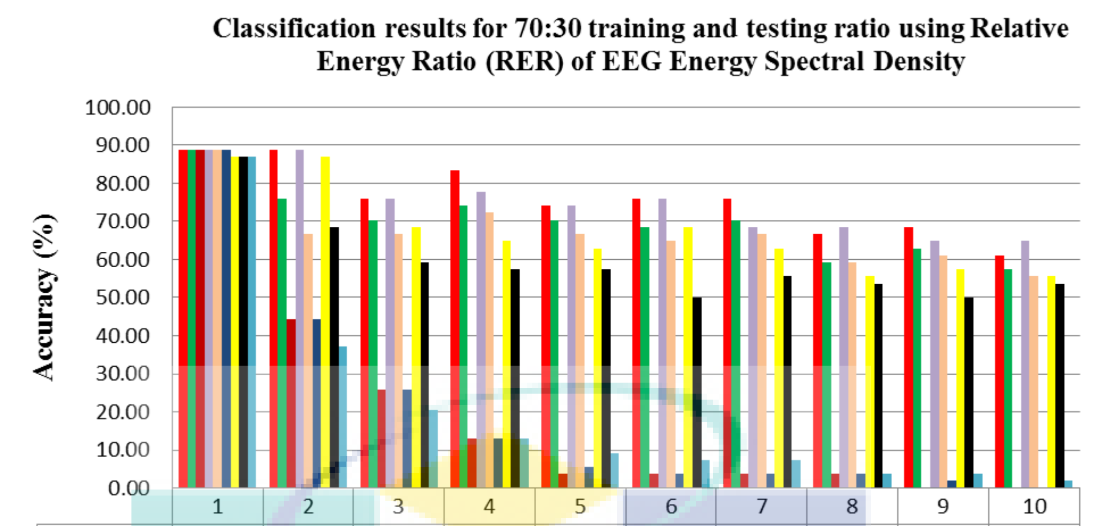


	1	2	3	4	5	6	7	8	9	10
Euc&Near	72.22	72.22	72.22	66.67	62.22	64.44	63.33	63.33	63.33	61.11
Euc&Rand	72.22	58.89	63.33	63.33	60.00	53.33	55.56	61.11	62.22	56.67
Euc&Con	72.22	35.56	22.22	13.33	8.89	3.33	2.22	1.11	0.00	0.00
City&Near	74.44	74.44	71.11	68.89	65.56	66.67	67.78	67.78	63.33	60.00
City&Rand	74.44	54.44	62.22	66.67	56.67	60.00	63.33	61.11	60.00	55.56
City&Con	74.44	33.33	20.00	15.56	11.11	7.78	4.44	2.22	1.11	0.00
Cos&Near	68.89	68.89	60.00	57.78	56.67	58.89	55.56	56.67	55.56	55.56
Cos&Rand	68.89	55.56	53.33	57.78	52.22	47.78	47.78	54.44	51.11	53.33
Cos&Con	68.89	28.89	18.89	12.22	7.78	5.56	5.56	2.22	1.11	0.00

Figure 5.36: The Classification Results using Relative Energy Ratio (RER) of ESD at 50:50 Training and Testing ratio

5.4.2.2 70:30 Training to Testing Ratio

Another experiment was conducted to classifiy the experimental group using relative energy ratio as an input to classifier, and this time the data is split into 70:30 training to testing ratio while feeding the same features to the classifier. The classification results shown in Figure 5.37 shows that the classification accuracy has increased from 74.44% to 88.89%, which is an outstanding improvement in the classification process. The classification accuracy is improved by 14.5%. Again, besides the selected input features, the best classification results are obtained when the classifier are configured to distance of Euclidean and Cityblock, and rule of nearest, random and consensus with smaller value of k (k=1).



	1	2	3	4	5	6	7	8	9	10
Euc&Near	88.89	88.89	75.93	83.33	74.07	75.93	75.93	66.67	68.52	61.11
Euc&Rand	88.89	75.93	70.37	74.07	70.37	68.52	70.37	59.26	62.96	57.41
Euc&Con	88.89	44.44	25.93	12.96	3.70	3.70	3.70	3.70	0.00	0.00
City&Near	88.89	88.89	75.93	77.78	74.07	75.93	68.52	68.52	64.81	64.81
City&Rand	88.89	66.67	66.67	72.22	66.67	64.81	66.67	59.26	61.11	55.56
City&Con	88.89	44.44	25.93	12.96	5.56	3.70	3.70	3.70	1.85	0.00
Cos&Near	87.04	87.04	68.52	64.81	62.96	68.52	62.96	55.56	57.41	55.56
Cos&Rand	87.04	68.52	59.26	57.41	57.41	50.00	55.56	53.70	50.00	53.70
Cos&Con	87.04	37.04	20.37	12.96	9.26	7.41	7.41	3.70	3.70	1.85

Figure 5.37: The Classification Results using Relative Energy Ratio (RER) of ESD at 70:30 Training and Testing ratio

Some observations regarding the use of k-NN to classify EEG features are as follows. The k value can be varied according to the size of the data, however, the highest classification rate is achieved when k is small. The distance and rule of the classifier also affect the classification accuracy, and it is observed that the highest classification rate is frequently observed at a distance of Euclidean and rule of the nearest. The smallest of k (< 10), a distance of Euclidean and rule of nearest is the best combination of k-NN classifier setting to produce better accuracy. Besides the features and classifier setting, other parameters that affect the classifier performance is the testing to training ratio of the data, where it is noted that the best ratio to use during the classification process is 70:30.

From the experimental results, the highest classification rate achieved is 88.89% using RER of energy spectral density as a feature to the classifier with 70:30 training to testing ratio. Meanwhile, for the classification configuration, the setting is optimum when Euclidean distance and nearest rule are used.

5.4.3 Summary of the Classification Results

The summary of the classification results are tabulated in Table 5.6 and Table 5.7. Based on the results, the energy ratio of EEG power spectrum is the best feature to detect human stress, in which case the classification accuracy increases from 74.44% to 88.89% when the training to testing ratio is varied from 50:50 to 70:30. Meanwhile, the highest accuracy is obtained when either k = 1 or k = 2. The classification results indicate that the best feature is RER and the best training and testing ratio is 70:30. Regardless of the features, the classification results using training and testing ratio at 70:30 is obviously outperform the classification results using 50:50 testing and training ratio. Therefore, the selection of the best EEG features and the best training and testing ratio with the optimal classification configuration has produced high classification accuracy.

Table 5.6:
Summary of Classification Accuracy using Different Features at 50:50 Training to Testing ratio

No.	Features	Formula	Ratios	Accuracy (%)	k
1	Normalization of Asymmetry	(R-L) / (R+L)	50:50	65.56	1,2
2	Natural log of Asymmetry	(In R – ln L)	50:50	66.67	1,2,7
3	Ratio of Energy of each frequency bands	ESD (δ, θ, α, β) / Total Power Spectrum	50:50	74.44	1,2

Table 5.7:
Summary of Classification Accuracy using Different Features at 70:30 Training to Testing ratio

No.	Features	Formula	Ratios	Accuracy (%)	k
1	Normalization of Asymmetry	(R-L) / (R+L)	70:30	68.52	1
2	Natural log of Asymmetry	(In R – ln L)	70:30	74.07	1,2
3	Ratio of energy of each frequency bands	ESD (δ, θ, α, β) / Total Power Spectrum	70:30	88.89	1,2

In the classification process, spectral centroids act as a target or class. Thus, the spectral centroids across the EEG frequency bands and for all experimental groups need to be calculated. The values of the spectral for all features are shown in Table 5.8. It is noted that the centroids value for Group 2 is highest at **1.72** when normalisation of Asymmetry of ESD is used as a feature, while the lowest value (**1.23**) is also achieved in Group 2 when RER is used as the feature.

Table 5.8:
Spectral Centroids of the Group based on the Selected EEG Features

Group	Normalization of Asymmetry of ESD	Natural Log Asymmetry of ESD	Relative Energy Ratio of ESD
1	1.47	1.30	1.53
2	**1.72**	1.25	**1.23**
3	1.62	1.25	1.49
4	1.46	1.30	1.46

The k-NN produces high classification accuracy using power ratio of ESD and training to testing ratio of 70:30, while spectral centroids are set as class for the classification process. In order to support the classifier performance, the classifier is also tested for sensitivity and specificity, and the classification results obtained when k is varied from 1 to 10 are shown in Table 5.9 while the rest of the results are shown in Appendix C.

Table 5.9:
The Summary of the Classification Accuracy, Sensitivity and Specificity at 70:30 Training to Testing ratio

k	**Accuracy (%)**	**Sensitivity (%)**	**Specificity (%)**
1	88.89	89.09	91.61
2	88.89	90.00	91.67
3	75.93	89.23	90.53

4	83.33	90.00	89.56
5	74.07	89.33	90.26
6	75.93	88.75	90.38
7	75.93	89.41	90.50
8	66.67	88.89	91.03
9	68.52	88.42	90.28
10	61.11	88.00	90.77

From the table, it is observed that the percentage of sensitivity and specificity remain high even when k is varied from 1 to 126, which is the maximum number of training data, while 54 of the data are the testing data. However, the classification results when k is varied from 1 to 10 is discussed in this thesis. In order to calculate the overall accuracy, sensitivity and specificity of the classification, the confusion matrix is employed as illustrated in Table 5.10. For this study, the testing data are tested for accuracy, sensitivity and specificity. Beside the classification accuracy of 88.89%, the classification sensitivity and specificity of 100% is observed in Group 4 and Group 2 respectively.

Table 5.10:
Results of Confusion Matrix

		Actual Class			
		1	2	3	4
Target Class	1	13	1	1	0
	2	0	12	0	0
	3	0	0	11	1
	4	0	3	0	12
Accuracy		88.89	88.89	88.89	88.89
Sensitivity		80.0	86.67	91.67	**100**
Specificity		97.23	**100**	97.37	90

Besides the overall classification performance in term of accuracy, sensitivity and specificity, the classification performance by group are also measured as depicted in Table 5.11. The classification performance of all groups are good. The overall classification accuracy for all groups is 88.89%. High sensitivity is observed in Group 2 while high specificity is noted in Group 4. As discussed in Chapter 3 (Section 3.6), sensitivity is defined as **true positive rate**. It indicates that Group 4 is correctly identified as non-stressed group. It also indicates that the intervention process (rotational bed) did affect the cognitive activities of the subjects in those group. Meanwhile, specificity is defined as **true negative rate**. It indicates that Group 2 is correctly identified as stressed group among the 3 non-stressed groups (Group 1, Group 3 and Group 4).

Table 5.11:
k-NN Classification Performance by Group

GROUP	2	4	3	1
ACCURACY	88.89%	88.89%	88.89%	88.89%
SENSITIVITY	86.67 %	**100%**	91.67%	80.0%
SPECIFICITY	**100%**	90%	97.37%	97.23%

5.5 CLASSIFICATION ERRORS

Figure 5.38 shows the plot of classification accuracy when k is varied from 1 to 10 for Asymmetry feature at 50:50 training to testing ratio. Meanwhile, the classification error in term of MSE is depicted by Figure 5.39. It is observed that at the highest classification accuracy rate (65.56%), the classification error is 0.09% (0.0009), and the total classification error is 1% (0.01) when the value of k is varied from 1 to 10. The classification error increases as the value of classifier neighborhood k increases, as can be observed in the rapid increase in the classification error when k = 8 and above.

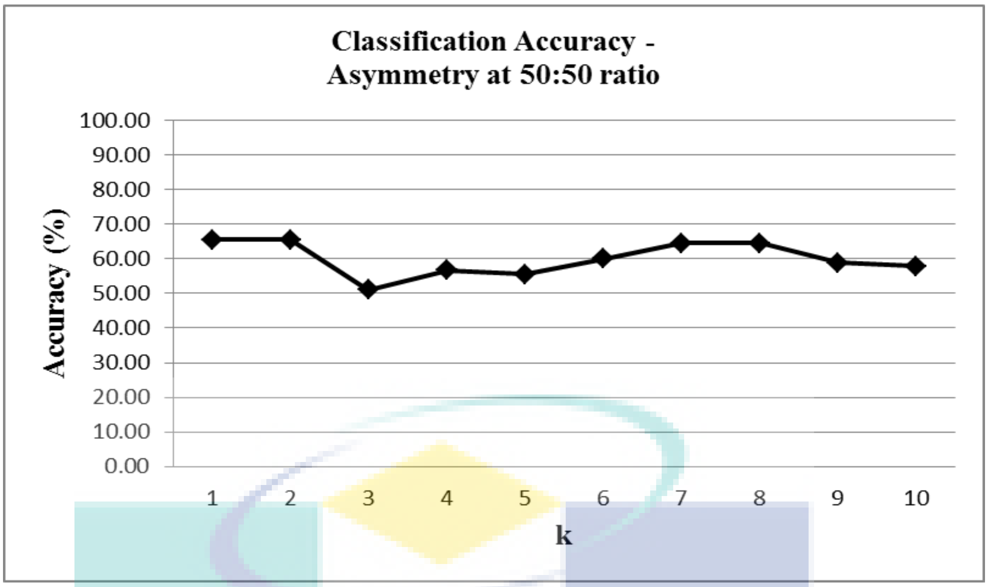


Figure 5.38: The Plot of Classification Accuracy Versus k for Asymmetry Feature at 50:50

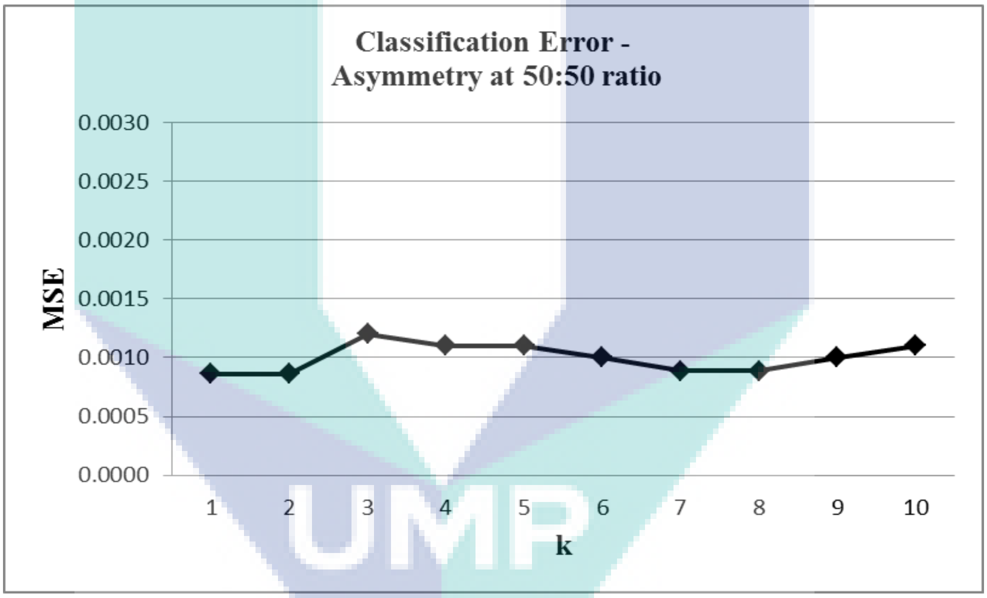


Figure 5.39: The Plot of Classification Error (MSE) Versus k for Asymmetry Feature at 50:50

Figure 5.40 shows the plot of classification accuracy when k is varied from 1 to 10 for Asymmetry feature at 70:30 training to testing ratio. Meanwhile, the classification error is shown in Figure 5.41. It is observed that at the highest classification accuracy rate (68.52%), the classification error is 1.33% (0.0133) and the total classification error is 18% (0.01) when the value of k is varied from 1 to 10. The classification error starts to increase as the value of classifier neighborhood k increases, as can be observed in the rapid increase in the classification error when k = 3 and above. Meanwhile, it is a rapid decrease of classification accuracy when the value of k increases. It is also observed that the classification error at 70:30 training to testing ratio is higher than classification error at 50:50 trainig to testing ratio.

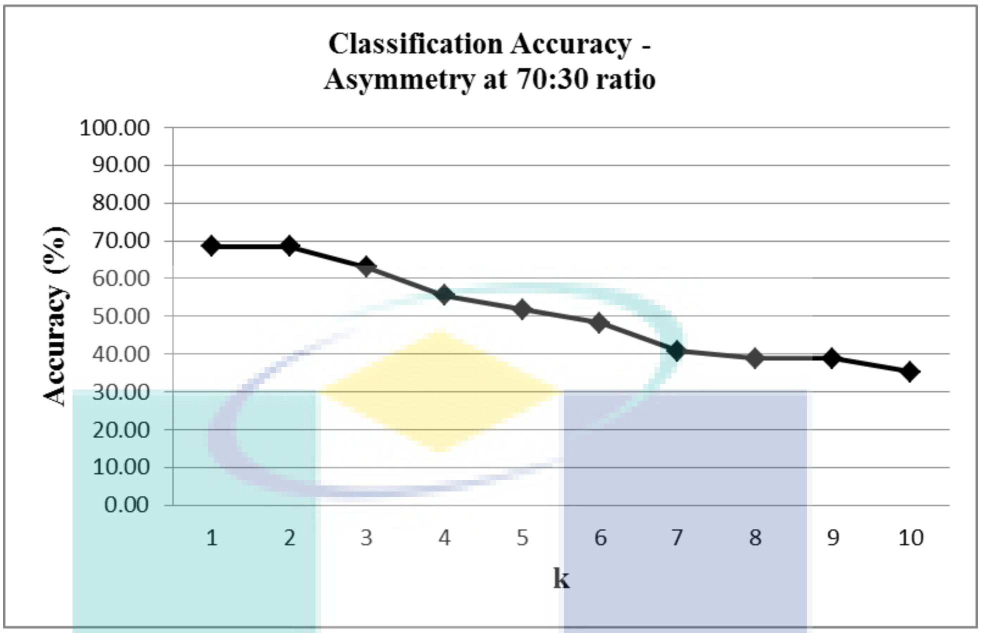


Figure 5.40: The Plot of Classification Accuracy Versus k for Asymmetry Feature at 70:30

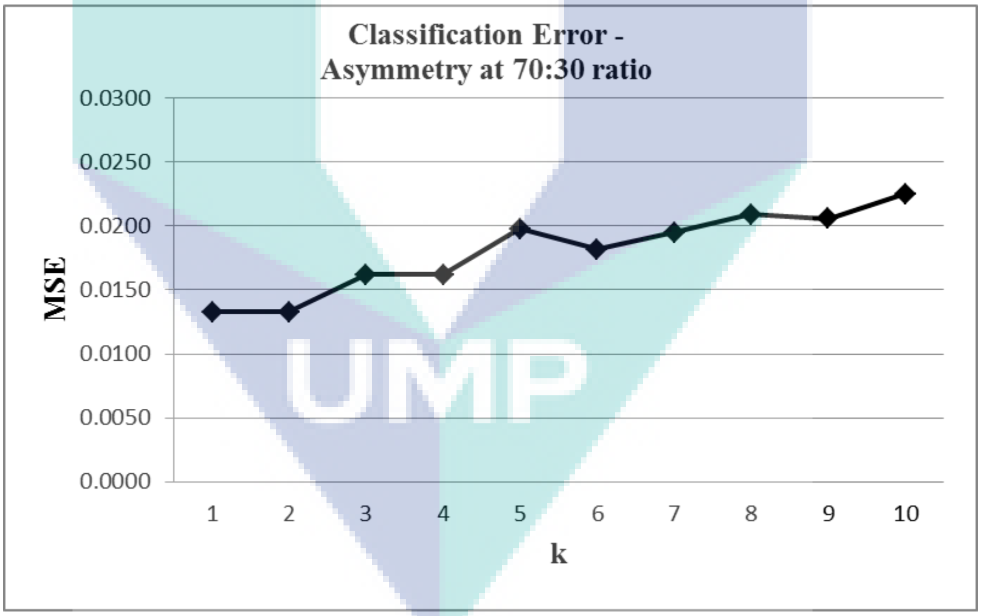


Figure 5.41: The Plot of Classification Error (MSE) Versus k for Asymmetry Feature at 70:30

Figure 5.42 shows the plot of classification accuracy when k is varied from 1 to 10 for Natural Log Asymmetry feature at 50:50 training to testing ratio. Meanwhile, the classification error is shown in Figure 5.43. It is observed that at the highest classification accuracy rate (66.67%), the classification error is 0.08% (0.0008) and the total classification error is 1% (0.01) when the value of k is varied from 1 to 10.

The classification error increases at k = 3 before decreasing until k = 7. Then, the classification error increases again when k above 7, it starts to increase as the value of classifier neighborhood k increases, as can be observed in the rapid increase in the classification error when k = 3 and above. Meanwhile, it is a rapid decrease of classification accuracy when the value of k increases. Regardless of the feature, it is also observed that the classification error is almost similar to the classification error produced by classification of Asymmetry feature at 50:50 training and testing ratio.

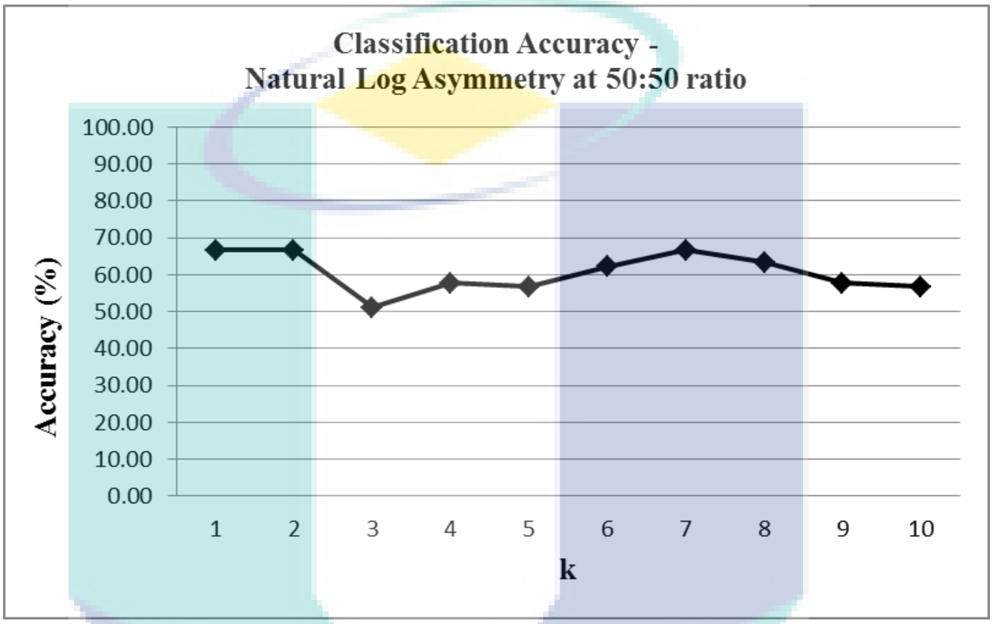


Figure 5.42: The Plot of Classification Accuracy Versus k for Natural Log Asymmetry Feature at 50:50

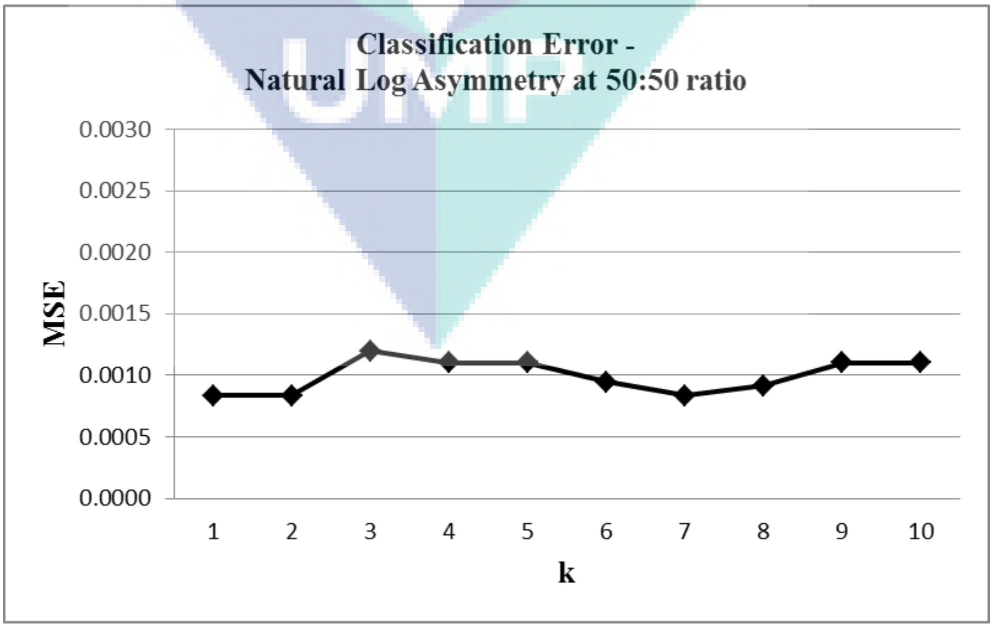


Figure 5.43: The Plot of Classification Error (MSE) Versus k for Natural Log Asymmetry Feature at 50:50

Figure 5.44 shows the plot of classification accuracy when k is varied from 1 to 10 for Natural Log Asymmetry feature at 70:30 training to testing ratio. Meanwhile, the classification error is shown in Figure 5.45. It is observed that at the highest classification accuracy rate (74.07%), the classification error is 0.06% (0.0006) and the total classification error is 0.8% (0.0085) when the value of k is varied from 1 to 10. It is a normal trend of the classification error until k reach 9. However, the classification accuracy increase to 70.37% at k = 10 which result the deacrease of classification error to 0.0007.

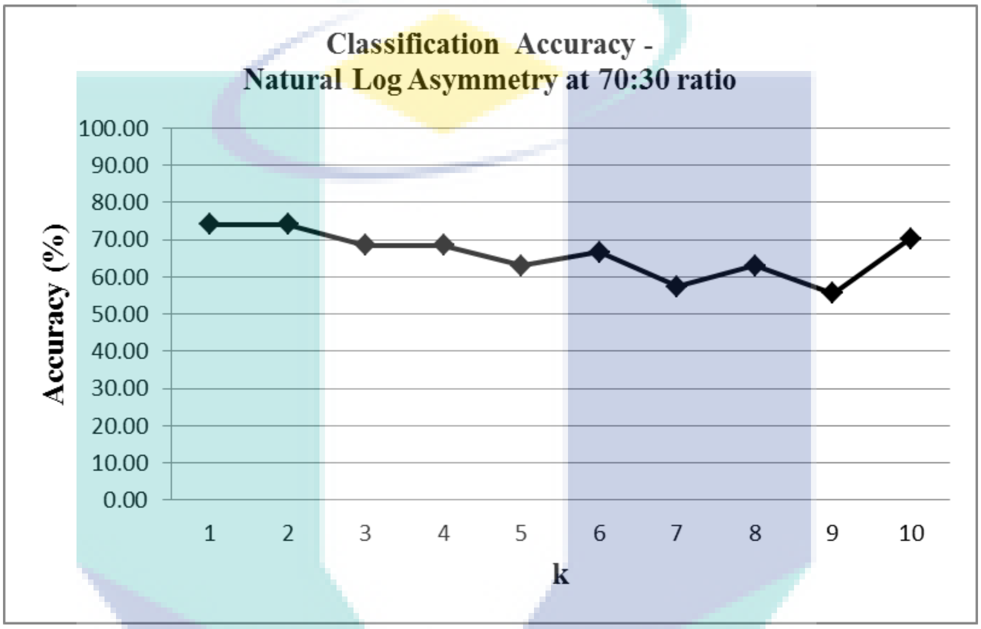


Figure 5.44: The Plot of Classification Accuracy Versus k for Natural Log Asymmetry Feature at 70:30

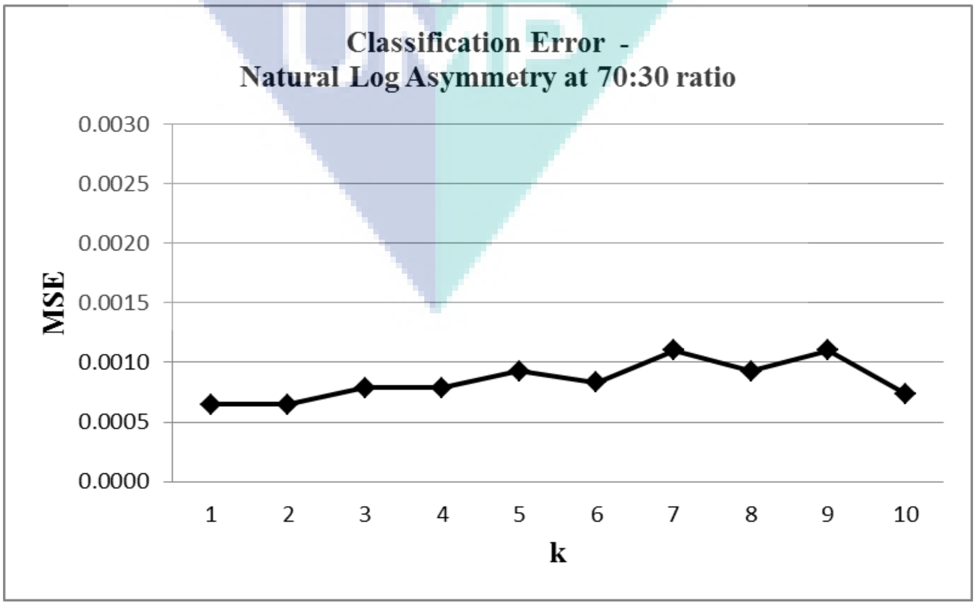


Figure 5.45: The Plot of Classification Error (MSE) Versus k for Natural Log Asymmetry Feature at 70:30

Figure 5.46 shows the plot of classification accuracy when k is varied from 1 to 10 for RER feature at 50:50 training to testing ratio. Meanwhile, the classification error is shown in Figure 5.47. It is observed that at the highest classification accuracy rate (74.44%), the classification error is 0.4% (0.0042) and the total classification error is 6% (0.0612) when the value of k is varied from 1 to 10. There is a normal trend for both classification accuracy and error except for the classification error at k = 5, 6 and 7. At these values of k, the classification error suppose to be in increasing trend. It might contributes to the high classification error and affect the classification accuracy at those value of k. The classification accuracy decreases when the value of k increases. The highest classification always obtained at the lowest value of k.

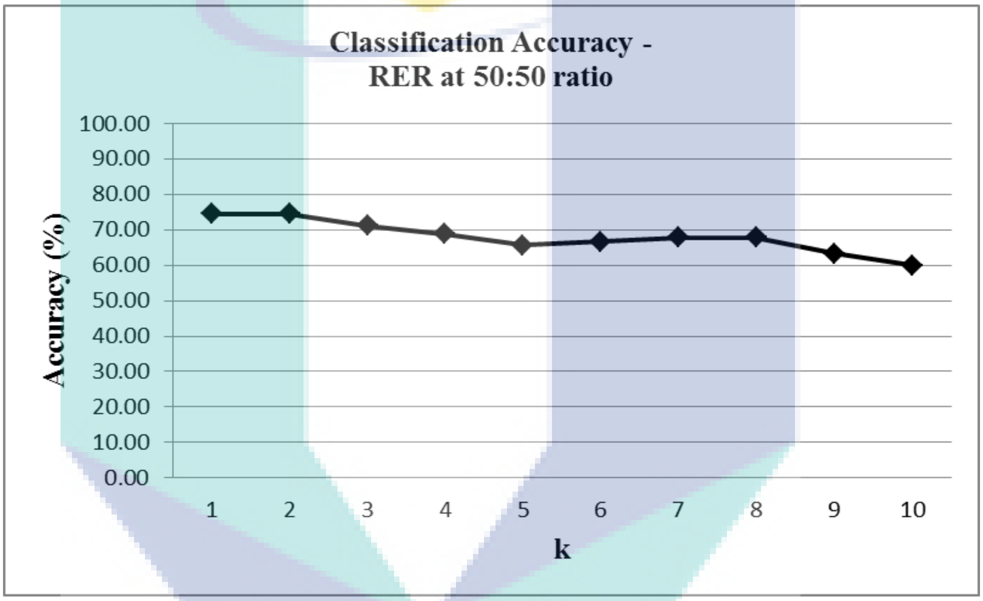


Figure 5.46: The Plot of Classification Accuracy Versus k for RER Feature at 50:50

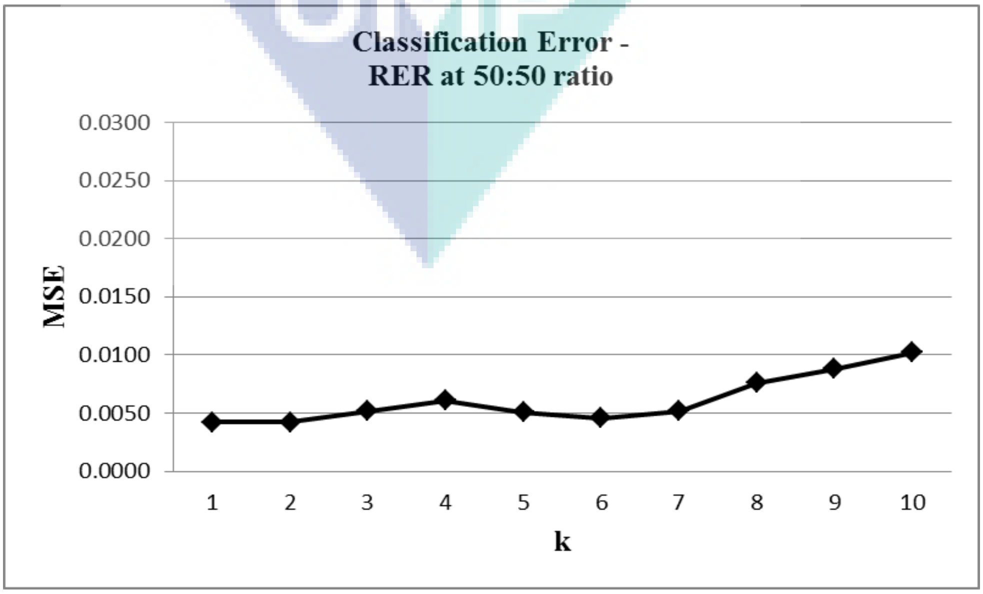


Figure 5.47: The Plot of Classification Error (MSE) Versus k for RER Feature at 50:50

Figure 5.48 shows the plot of classification accuracy when k is varied from 1 to 10 for RER feature at 70:30 training to testing ratio. Meanwhile, the classification error is shown in Figure 5.49. It is observed that at the highest classification accuracy rate (88.89%), the classification error is 0.25% (0.0025) and the total classification error is 3.9% (0.0386) when the value of k is varied from 1 to 10. There is a normal trend for both classification accuracy and error. The classification error increases as the value of classifier neighborhood k increases, as shown by the rapid increase in the error when k = 8 and above, which tally with the findings from previous studies where the highest classification rate is achieved at the small value of k. At the highest classification accuracy, the classification error of 0.0025 can be considered minor.

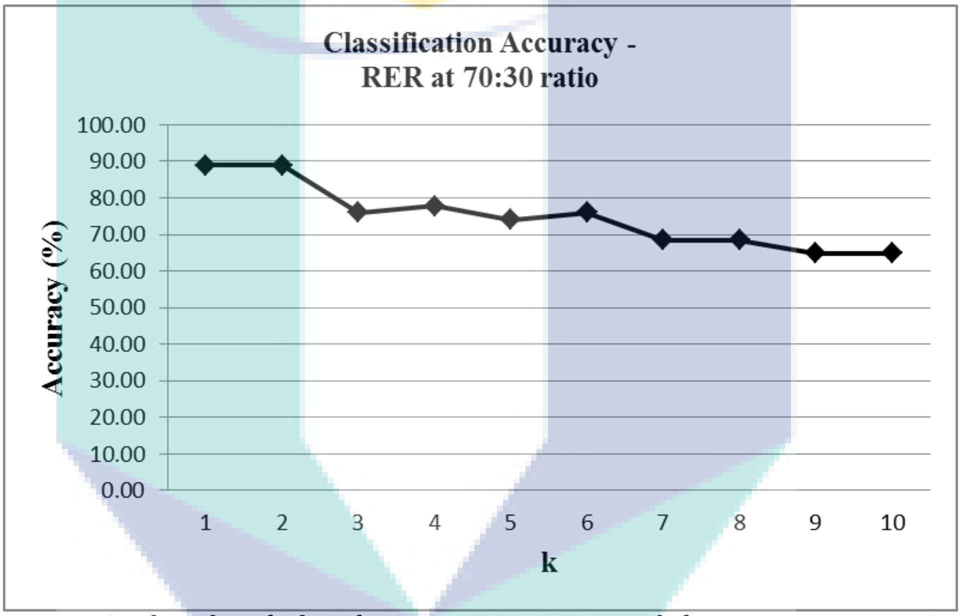


Figure 5.48: The Plot of Classification Accuracy Versus k for RER Feature at 70:30

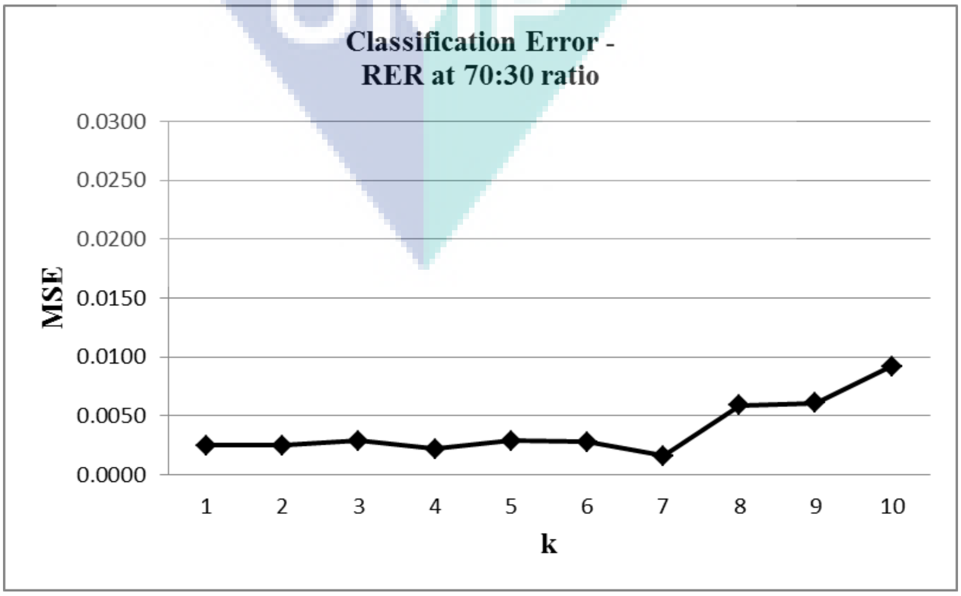


Figure 5.49: The Plot of Classification Error (MSE) Versus k for RER Feature at 70:30

5.6 RESULTS OF CROSS-VALIDATING THE CLASSIFIER

This section will verify the robustness of the classifier by cross-validating the classifier using k-fold cross-validation and leave-one-out techniques, as previously discussed. The results of 10-fold cross-validation and leave-one out cross-validation are shown in Figure 5.50. The 10-fold cross-validation yields an overall accuracy of 78.89%, which is slightly lower than that of k-NN classification and leave-one cross-validation. However, its performance can be considered satisfactory as the input data to k-NN classifier, regardless of the size of the training and testing data, are divided into 10 portions before being validated. Furthermore, its classification performance in terms of sensitivity and specificity at 92% and 91% are good. Meanwhile, the overall classification accuracy of cross-validating k-NN classifier using leave-one-out technique is approximately 5% lower than that obtained from the normal classification technique using different training to testing ratios. However, the leave-one-out cross-validation technique produces 100% sensitivity and specificity. Therefore, from the outstanding k-NN classifier performance obtained in terms of overall accuracy, sensitivity and specificity, it can be concluded that the classifier is robust as the classification results are quite precise.

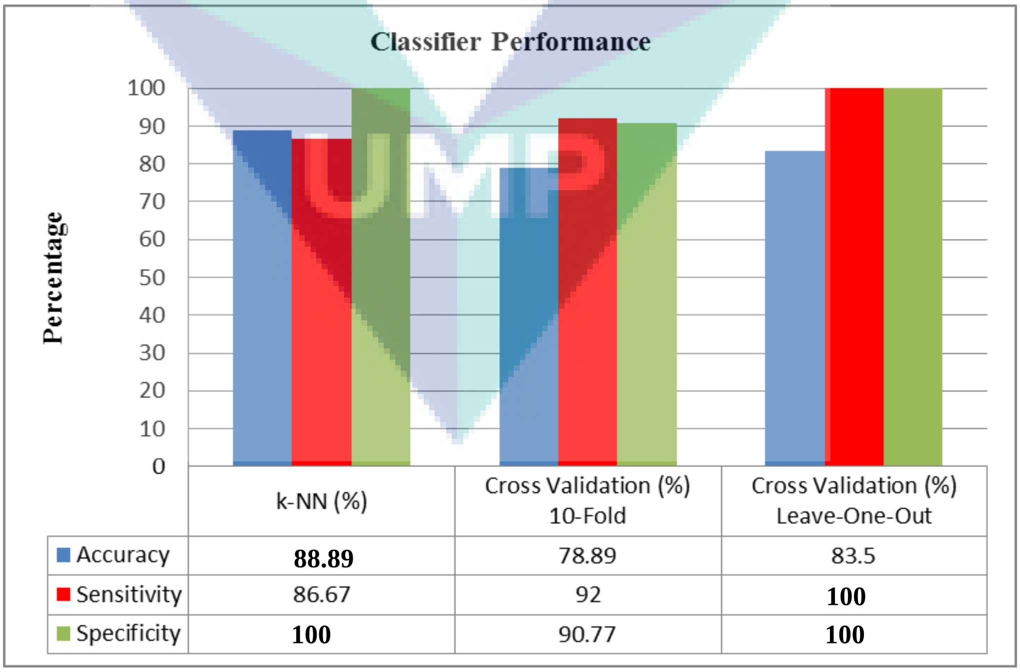


Figure 5.50: The Performance of k-NN Classifier and Cross-Validation

5.7 CLUSTERING PERFORMANCE

Clustering of the groups involved in this study using FCM and FKM is implemented to support the classification results obtained using k-NN in trying to determine the stress group. The results of FCM clustering are shown in Figure 5.51, where it is noted that the 180 RER data of EEG energy spectral densities are clearly clustered into three groups. The data are clustered according to the EEG frequency bands, which are Beta band cluster, Alpha band cluster and Theta band cluster. However, the centroids of Theta band and Delta band overlap each other, which results in three different clusters. Cluster 1 and cluster 2 indicate RER data at Delta and Theta band. Meanwhile, Cluster 3 and Cluster 4 represent RER at Beta band and Alpha band, respectively. In addition, it is observed that the data in Cluster 3 are more widely distributed compared to that of Cluster 4. Hence, FCM is capable of identifying the group that might have stress features. The clustering performance achieved is 80% and is calculated using the index of the maximum membership (U) for each cluster, where the spectral centroids are applied as a class in classification process. Also, the centroids value from four features (RER of Delta band, Theta band, Alpha band and Beta band) from the four groups involved in the study are clustered accordingly.

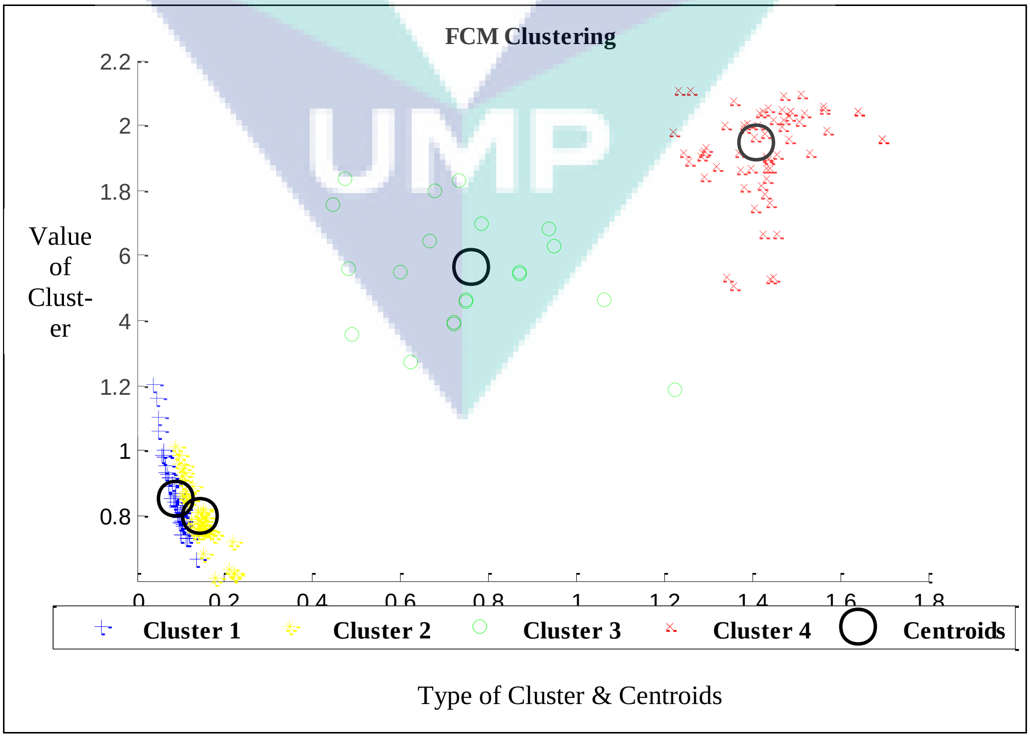


Figure 5.51: FCM Clustering Results

The results of clustering RER from each group and frequency bands using FKM is illustrated by Figure 5.52. Unlike FCM clustering, FKM clustering is able to produce four prominent clusters. It can be seen that the data distribution in the new clustering is less than that of the original clustering as the data that was far away from the centroids of the cluster center are removed. The partition of the datasets containing the EEG features (RER) using FKM has improved the clustering accuracy by 3%. The clustering performance are approximately the same as that of classification accuracy using k-NN classifier.

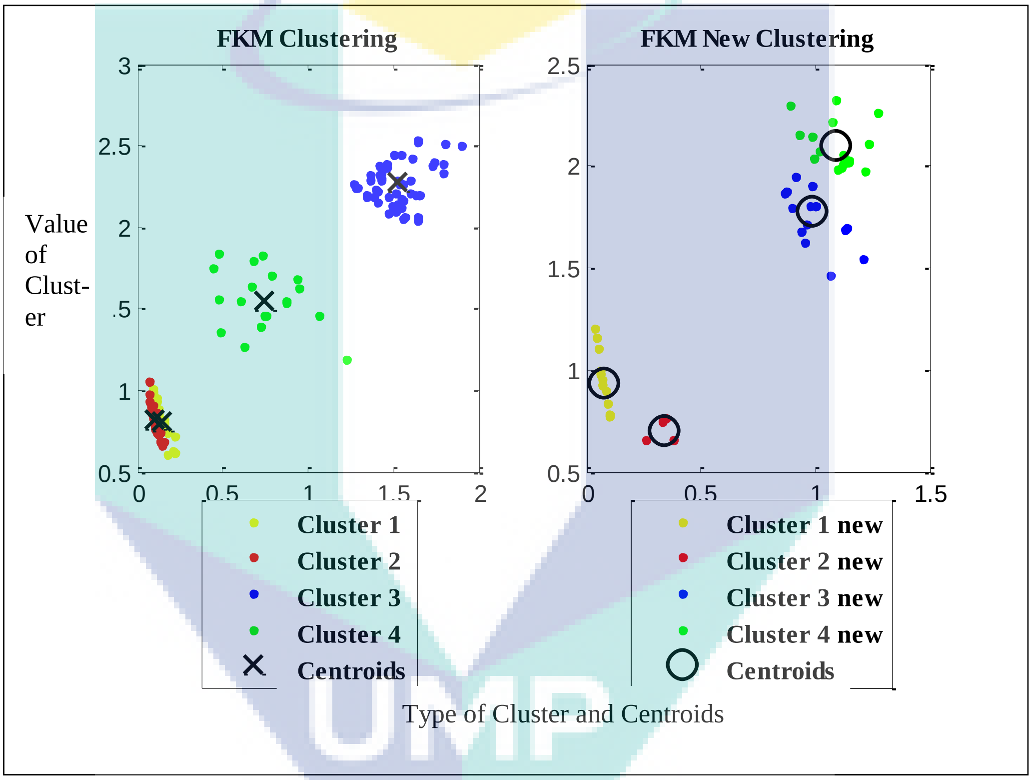


Figure 5.52: FKM Clustering Results

5.8 RESULTS OF STATISTICAL ANALYSES

In order to support the results of the classification and clustering processes, the RER of the EEG power spectrum that has been selected as the EEG feature, are statistically tested for the significant value. Thus, 180 datasets containing the features are analyzed using ANOVA and the results of applying ANOVA to the EEG features (RER, Asymmetry ratio and Natural Log Asymmetry ratio) are shown in Table 5.12, Table 5.13 and Table 5.14. Referring to Table 5.12, the ratio of Alpha energy to Theta energy are significant between groups and within groups since their *p*-value is lower than 0.05. The results of ANOVA for Alpha and Theta energy ratios is $F(3, 176) = 8.62$, $p = 0.00$ and $F(3, 176) = 9.48$, $p = 0.00$, respectively. However, the ratio of Beta energy to Delta energy between groups and within groups are not significant since their *p*-values are greater than 0.05. The statistical results given by ANOVA match with the data distribution in FCM and FKM.

Table 5.12:
ANOVA Results for RER Features across the EEG Frequency Bands

Features		Sum of Squares	df	Mean Square	F	Sig.
Beta Ratio	Between Groups	.205	3	.068	.919	.433
	Within Groups	13.117	176	.075		
	Total	13.323	179			
Alpha Ratio	Between Groups	2.095	3	.698	8.617	**.000**
	Within Groups	14.263	176	.081		
	Total	16.358	179			
Theta Ratio	Between Groups	.243	3	.081	9.484	**.000**
	Within Groups	1.501	176	.009		
	Total	1.743	179			
Delta Ratio	Between Groups	.023	3	.008	2.438	.066
	Within Groups	.558	176	.003		
	Total	.581	179			

The results obtained after applying ANOVA to EEG asymmetry and natural log asymmetry ratios are shown in Table 5.13 and Table 5.14, respectively. There are no statistically significant difference observed on asymmetry and natural log asymmetry between groups and within groups since their *p*-value is greater than 0.05. For asymmetry ratio, the ANOVA results for Beta band and Alpha band are $F(3, 176) = 0.68$, $p = 0.56$ and $F(3, 176) = 1.06$, $p = 0.37$, respectively. Also, the F and *p* value for the Beta band and Alpha band of natural log asymmetry ratio are almost the same as that of the asymmetry ratio. The results of ANOVA indicate that the energy ratios of EEG frequency bands are the preferred EEG features in order to produce good classification and clustering in determining the stress group.

Table 5.13:
ANOVA Results for Asymmetry Ratio across the EEG Frequency Bands

Features		Sum of Squares	df	Mean Square	F	Sig.
Beta_Asym	Between Groups	.055	3	.018	.681	.565
	Within Groups	4.764	176	.027		
	Total	4.819	179			
Alpha_Asym	Between Groups	.044	3	.015	1.063	.366
	Within Groups	2.437	176	.014		
	Total	2.481	179			
Theta_Asym	Between Groups	.081	3	.027	1.811	.147
	Within Groups	2.612	176	.015		
	Total	2.693	179			
Delta_Asym	Between Groups	.077	3	.026	1.013	.388
	Within Groups	4.458	176	.025		
	Total	4.535	179			

Table 5.14:
ANOVA Results for Natural Log Asymmetry Ratio across the EEG Frequency Bands

Features		Sum of Squares	df	Mean Square	F	Sig.
Beta_ln_Asym	Between Groups	.212	3	.071	.681	.565
	Within Groups	18.259	176	.104		
	Total	18.471	179			
Alpha_ln_Asym	Between Groups	.190	3	.063	1.086	.357
	Within Groups	10.285	176	.058		
	Total	10.475	179			
Theta_ln_Asym	Between Groups	.343	3	.114	1.843	.141
	Within Groups	10.921	176	.062		
	Total	11.264	179			
Delta_ln_Asym	Between Groups	.315	3	.105	.978	.404
	Within Groups	18.871	176	.107		
	Total	19.186	179			

Besides ANOVA, the Pearson correlation study or scatterplot is conducted in order to investigate the relationship between the change of the cognitive states (eyes-closed and eyes-open) with energy spectral density of the EEG frequency bands in the experimental groups. Table 5.15 – 5.17 and Figure 5.53 – 5.55 illustrate the linear relationship between EEG features and cognitive state (experimental groups). It is observed that there is a linear correlation between the EEG features with the cognitive state, EC and EO states (represented by Group 1 and Group 2 respectively). For the RER feature, the good correlation is observed in Group 2 where the correlation strength for EEG Beta and Alpha band is **0.576** as illustrated by Table 5.15. Meanwhile, for the asymmetry ratio feature, the good correlation is also observed in Group 2 where the correlation strength is **0.629** as shown in Table 5.16. Next, for natural log asymmetry ratio, the good correlation is observed on Group 2 and Group 4 where the correlation strength of Beta and Alpha bands is **0.622** as depicted by Table 5.17. In term of correlation significant, the correlation results indicate the significant results where $p < 0.05$ for all EEG features and groups being studied.

Table 5.15:
Pearson Correlation Results using RER Feature Across EEG Frequency Bands

Features		Beta Ratio	Alpha Ratio	Theta Ratio	Delta Ratio
Beta Ratio	Pearson Correlation	1	**.576****	.340**	**-.564****
	Sig. (2-tailed)		.000	.000	.000
	N	180	180	180	180
Alpha Ratio	Pearson Correlation	**.576****	1	.466**	**-.871****
	Sig. (2-tailed)	.000		.000	.000
	N	180	180	180	180
Theta Ratio	Pearson Correlation	.340**	.466**	1	**-.651****
	Sig. (2-tailed)	.000	.000		.000
	N	180	180	180	180
Delta Ratio	Pearson Correlation	**-.564****	**-.871****	**-.651****	1
	Sig. (2-tailed)	.000	.000	.000	
	N	180	180	180	180

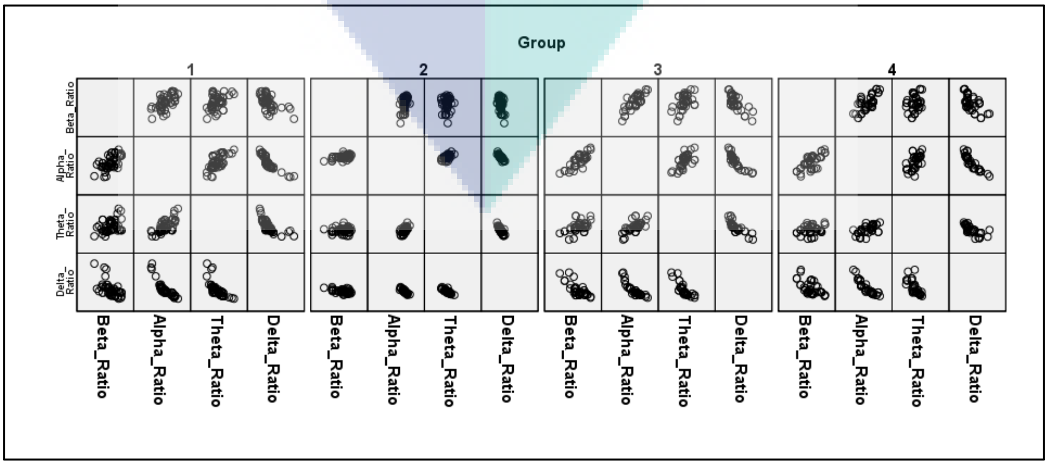


Figure 5.53: Scatterplot of the RER Feature from the Experimental Groups

Only the correlation between EEG Beta and Alpha bands from the 4 experimental groups are considered in this study since the changes on the EEG power spectrum of Beta and Alpha bands will indicate the existence of stress. Thus, the correlation strength and direction of Beta and Alpha bands in Group 1 (EC state) and Group 2 (EO state) will be the focus in this study. However, based on the correlation results, correlation strength between Group 1 and Group 2 using asymmetry and natural log asymmetry have outperformed the correlation strength using RER feature. The correlation strength between 0.5 to 1.0 is considered good.

Table 5.16:
Pearson Correlation Results using Asymmetry Feature Across EEG Frequency Bands

Features		Beta_Asym	Alpha_Asym	Theta_Asym	Delta_Asym
Beta_Asym	Pearson Correlation	1	**.629****	**.500****	.247**
	Sig. (2-tailed)		.000	.000	.001
	N	180	180	180	180
Alpha_Asym	Pearson Correlation	**.629****	1	**.804****	.357**
	Sig. (2-tailed)	.000		.000	.000
	N	180	180	180	180
Theta_Asym	Pearson Correlation	**.500****	**.804****	1	**.699****
	Sig. (2-tailed)	.000	.000		.000
	N	180	180	180	180
Delta_Asym	Pearson Correlation	.247**	.357**	**.699****	1
	Sig. (2-tailed)	.001	.000	.000	
	N	180	180	180	180

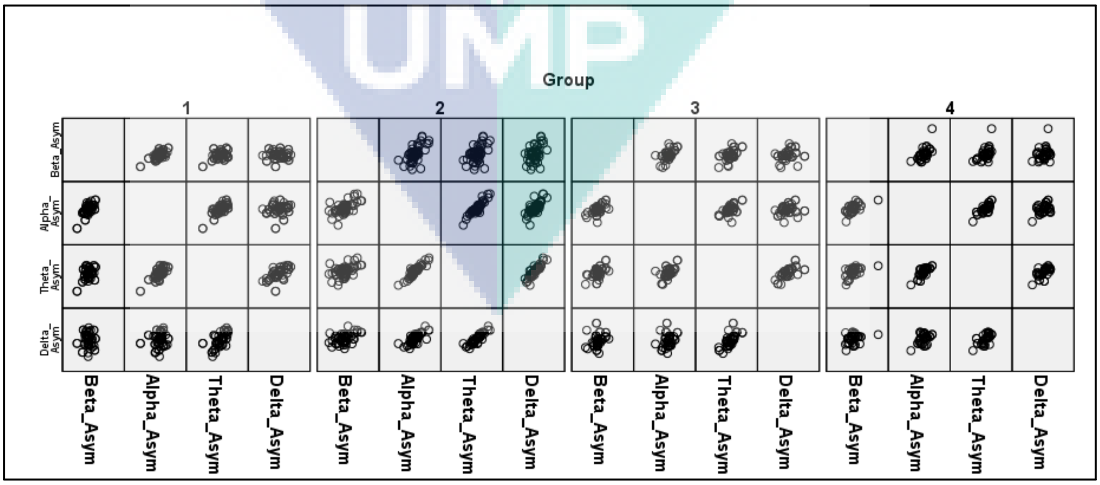


Figure 5.54: Scatterplot of the Asymmetry Feature from the Experimental Groups

Table 5.17:
Pearson Correlation Results using Natural Log Asymmetry Feature Across EEG Frequency Bands

Features		Beta_ ln_Asym	Alpha_ ln_Asym	Theta_ ln_Asym	Delta_ ln_Asym
Beta_ln_Asym	Pearson Correlation	1	**.622****	**.501****	.241**
	Sig. (2-tailed)		.000	.000	.001
	N	180	180	180	180
Alpha_ln_Asym	Pearson Correlation	**.622****	1	**.808****	.352**
	Sig. (2-tailed)	.000		.000	.000
	N	180	180	180	180
Theta_ln_Asym	Pearson Correlation	**.501****	**.808****	1	**.692****
	Sig. (2-tailed)	.000	.000		.000
	N	180	180	180	180
Delta_ln_Asym	Pearson Correlation	.241**	.352**	**.692****	1
	Sig. (2-tailed)	.001	.000	.000	
	N	180	180	180	180

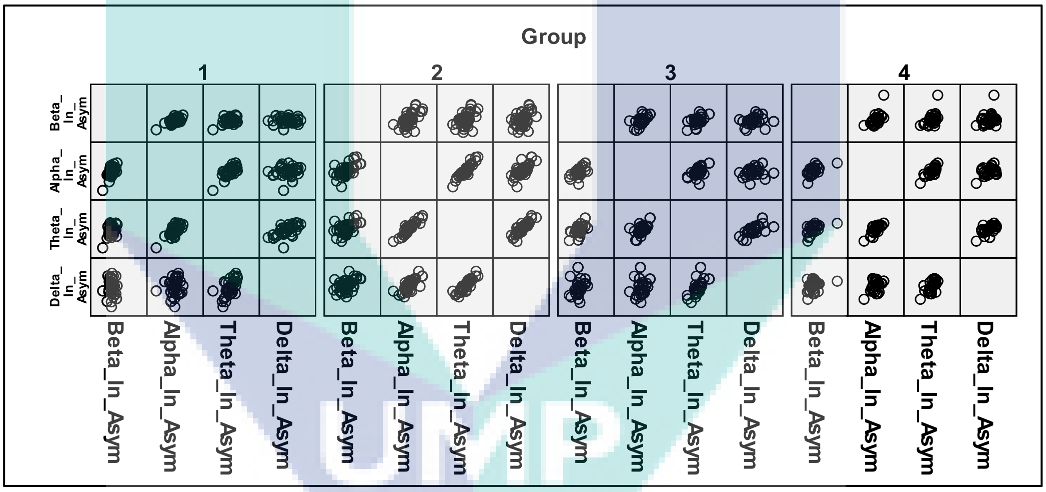


Figure 5.55: Scatterplot of the Natural Log Asymmetry Feature from the Experimental Groups

5.9 ASSIGNMENT OF STRESS INDEX

The elements of stress of any group involved in the study can be determined and consequently, the stress index as shown in Table 5.18 can be assigned by studying the EEG signals captured from the participants. The can be achieved based, on the results obtained after applying Shannon’s entropy on certain selected features of the raw EEG signals, finding the spectral centroids of the energy spectral densities of the signals, and finally, classifying and clustering the selected features. In the table is displayed only the results obtained after applying Spectral Centroids, Shannon Entropy, Average Energy of EEG Alpha band and Beta band across the groups, which is the focus of this study, while features from the Delta band and Theta band are not taken into account in developing the stress index. In this study, Index 1, which represents low stress, is assigned to Group 1 as its centroid and entropy are the highest among the groups, while its average energy of Beta band is the lowest. Index 2, which indicates moderate stress, is assigned to Group 3 and Group 4 as the groups’ average energy of Beta band is higher than that of Group1 but lower than that of Group 2, while the groups’ centroid of Alpha band is lower than that of Group 1 but higher than that of Group 2. Furthermore, the entropy of Group 3 and Group 4 is higher than that of Group 2. Consequently, Index 3, which indicate high stress, is assigned to Group 2.

However, the most noteworthy indicator of stress is the value of the average energy ratio of Alpha band and Beta band. It is noted that Group 2 has a much lower average energy ratio of Alpha band than that of the other groups, while its average energy ratio of the Beta band is higher than that of the other groups. These values and the assignment of the stress index match with the findings from previous studies on stress assessment using EEG signals [1, 33, 42, 68, 71, 136, 141, 144]. In those literature reviews, the major indicator for the existence of stress in human is based on the change in EEG Alpha and Beta energy when human encounters stress and the human body reacts to the stressors. Hence, it can be concluded that certain selected EEG features, the ratio of Average Energy and the spectral centroids in the Alpha-band and Beta-band, and Shannon Entropy can be used to recognize the level of stress from the four experimental groups involved in the study. The application of SC on RER of EEG frequency bands and the trend of SE are capable to detect the change in cognitive state from relax state to alert state due to the existence of the stressors.

Table 5.18:
Assignment of Stress Index

Group	Centroid (α)	Centroid (β)	Shannon Entropy (SE)	Average Energy (α)(%)	Average Energy (β)(%)	Stress Index
1	7.59	4.35	0.31	8.93	1.17	1
2	**3.07**	3.62	**0.29**	**3.79**	**1.38**	**3**
3	4.78	3.30	0.3	7.41	1.20	2
4	4.46	3.67	0.33	8.76	1.33	2

5.10 VERIFICATION OF STRESS INDEX

The assignment of the stress index stated in Table 5.18 need to be verified, and one method to verify the index is to use the Z-score technique, which is based on the average and standard deviation of the relative energy ratio of Alpha and Beta power in the group. The block diagram for stress index verification process using Z-score is illustrated in Figure 5.56. The EEG features is first selected before performing Z-transform to the features. Researchers have confirmed that Z-Score can be used to confirm the index using EEG features [200].

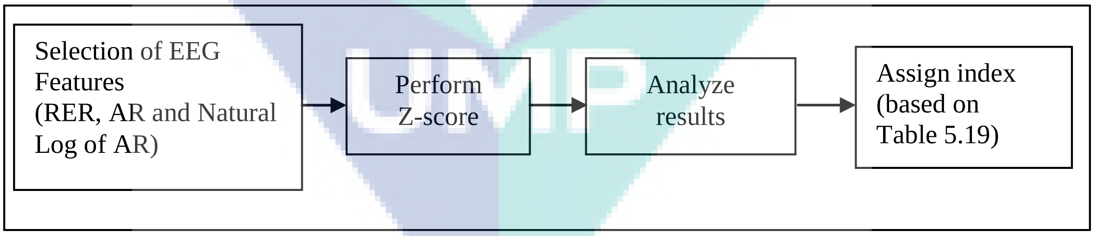


Figure 5.56: Block Diagram for Z-score Verification

The Z-score technique is only applied to the Beta and Alpha bands as the changes of the Beta and Alpha energy are the main indicator for the presence of stress. The Z-score used in this study is described in Table 5.19, where the index for the stress is assigned based on the location of the average energy ratio of the bands. Index 1 will be assigned when the Z-score is less than -1, Index 2 will be assigned when the score lies between -1 and +1, while Index 3 will be assigned when the score is greater than +1.

Table 5.19:
Z-score and Stress Index

Value of Z-score	Descriptions	Assigned Stress Index
less than -1	Low	1
greater than 1	High	3
between 1 and -1	Moderate	2

The average of the Beta and Alpha energy for each group involved in the experiments are calculated before implementing the Z-score on the mean energy, and the results obtained are shown in Table 5.20. Based on the Z-score of the Beta band, Index 3 was assigned to Group 2 as its Z-score is higher than 1, Index 2 was assigned Group 3 and Group 4 as both groups have a Z-score between -1 and 1, while Index 1 was assigned to Group 1 as its score is less than -1. As for the Alpha band, Index 1 is assigned to Group 2 as its Z-score is less than -1, while Index 2 is assigned to Groups 1, 3 and 4 as the Z-score of these groups is between -1 and +1, and no Index 3 was assigned to any of the groups. Theoretically, the index assigned to the Alpha band of Group 2 should be the greater than 1 and Index 3 would be assigned, which is the opposite that of the Beta band of Group 1 with stress Index 1. However, comparing the results shown in Table 5.18 and Table 5.20 indicates that the application of Z-score can be used to verify the assignment of stress index when Beta power across the group is used but not quite true when Alpha power is used.

Table 5.20:
Assignment of Stress Index using Z-score Technique

	Alpha			Beta		
Group	Power ratio	Z-score value	Stress index	Power ratio	Z-score value	Stress index
1	0.32083	0.82	2	0.227201	-1.317	1
2	**0.136149**	**-1.31**	**1**	**0.267834**	**1.030**	**3**
3	0.228489	-0.25	2	0.247518	-0.143	2
4	0.314533	0.74	2	0.257447	0.430	2

In order to figure out what happens during the implementation of Z-score on the power spectrum of EEG frequency bands across the groups (180 data), the values of Z-score of Alpha and Beta power are plotted as illustrated in Figure 5.57 and Figure 5.58, respectively. From the plots, it is observed that there is a no clear significant trends for the values of Z-score for Alpha power, except that the average of Alpha power from Group 2 (sample number 51 to sample number 100) is smaller than that of the other groups. However, there is a trend of high values of Z-score for Beta power from Group 2, which indicates high stress index. Here, the Z-score value for Beta power in Group 2 is much higher than that of the other groups. Hence, the results from the plot match with the assignment of stress index shown in Table 5.14. Figure 5.58 also indicates that the value of Z-score of Group 1 (sample number 1 to sample number 50) is smaller than that of Group 3 and Group 4 (sample number 101 to 180).

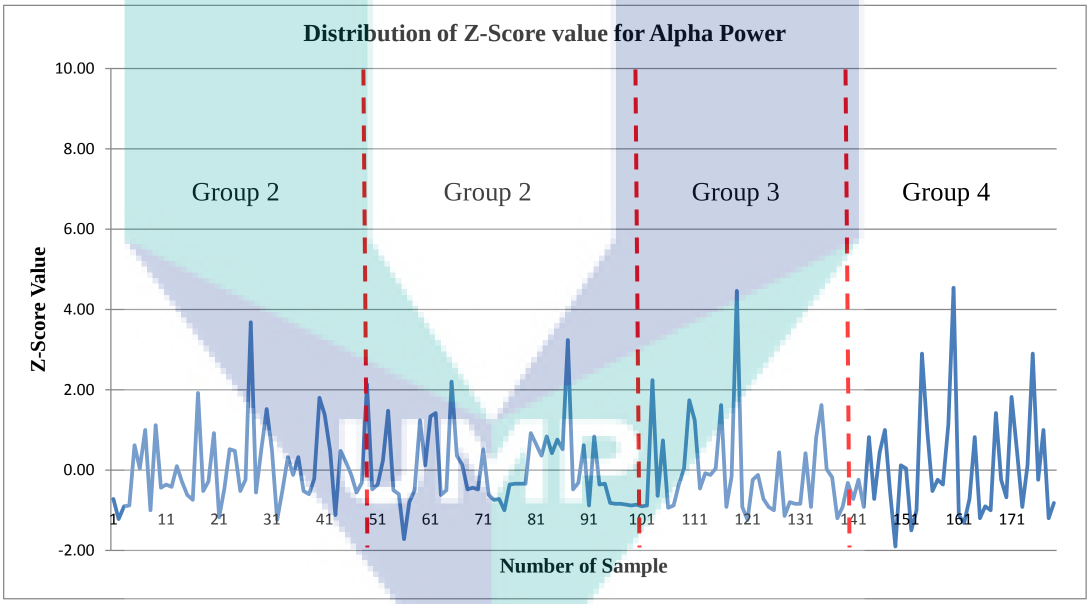


Figure 5.57: Distribution of Z-score value for Alpha Power

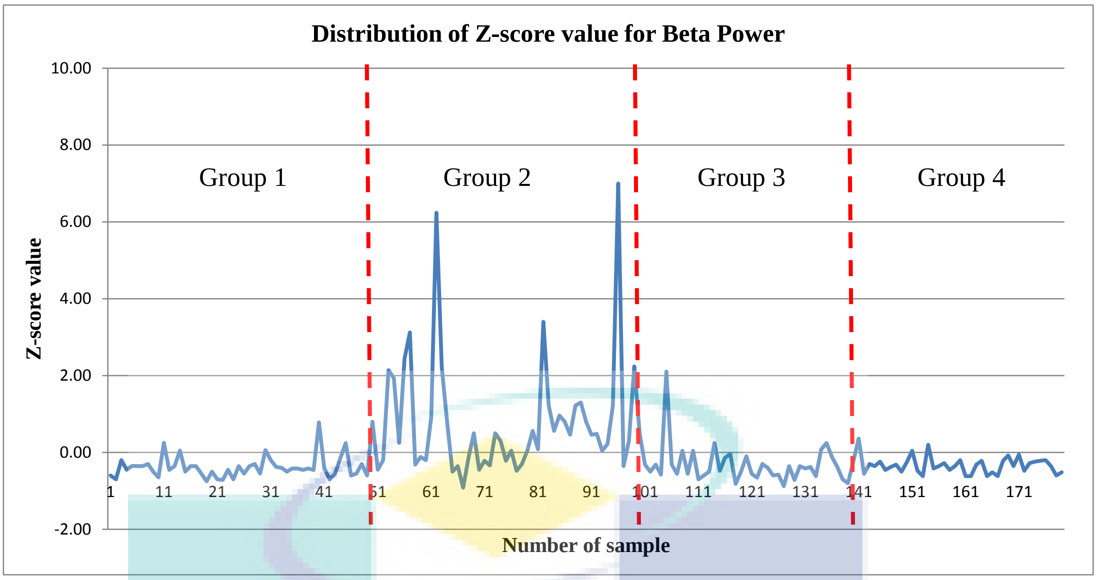


Figure 5.58: Distribution of Z-score value for Beta Power

Besides using Z-score, the verification of the stress index is implemented by checking the energy ratio of EEG Beta and Alpha bands of all subjects in the groups. In addition, the entropy of EEG data of each participant in the groups is also considered in the index verification process. First, the index is verified using the energy ratio of EEG data, where the energy ratio of the data within a specified range is computed using average and standard deviation of the data. The range of the data is represented by the two bold lines shown in Figure 5.59 to Figure 5.60. Since the average and standard deviation of the EEG data in the groups is different, the size of the range is also different. Figure 5.59 shows that the range of the Group 1 for Beta energy is from 0.40 to 2.00, where Index 1 (low stress) was assigned to the group. From this range, the number of subjects that belong to this index can be calculated. There are 9 data (Beta energy) lies outside the specified range, which means that 18% of the data in Group 1 do not belong to Index 1. However, only 4 EEG data (Alpha energy) in Group 1, or 8% of the data, do not belong to Index 1, as seen in Figure 5.60.

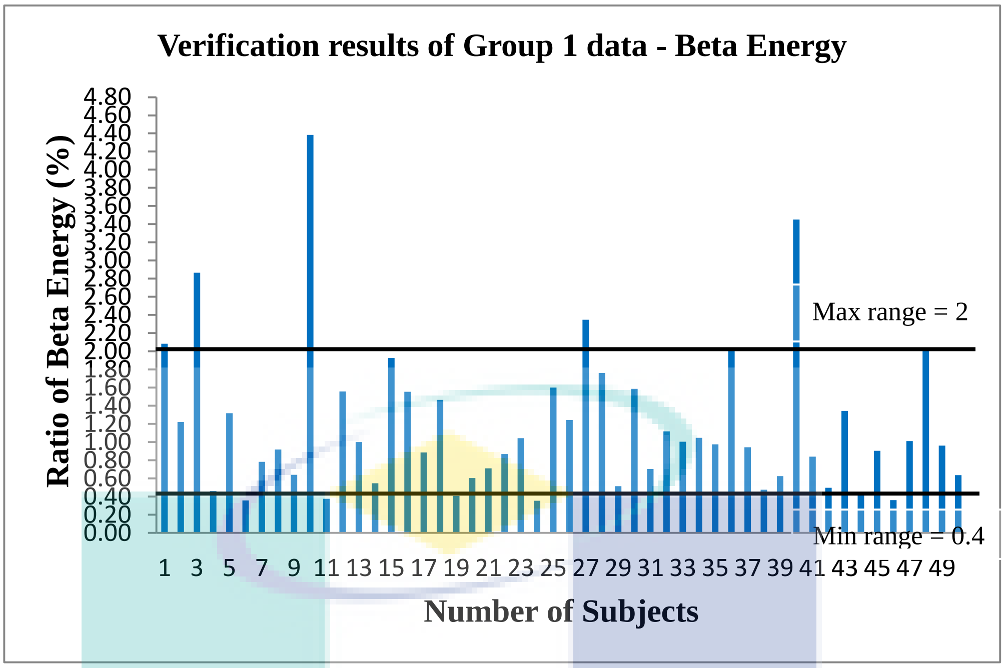


Figure 5.59: Verification Results for Stress Index 1 (Beta Energy) – Group 1

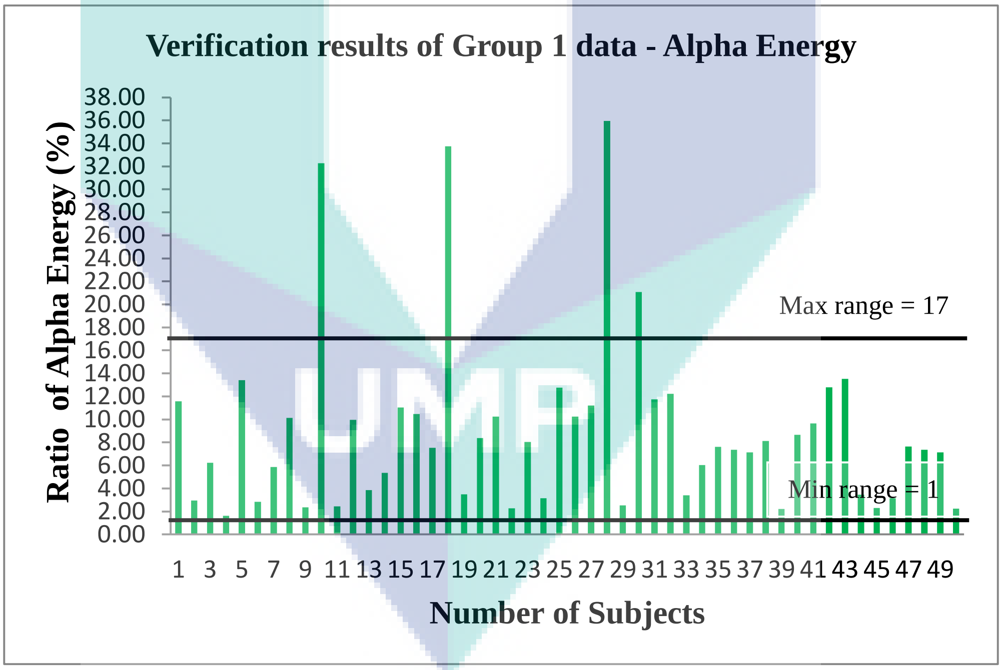


Figure 5.60: Verification Results for Stress Index 1 (Alpha Energy) – Group 1

For verification of Index 3, the highest level of stress, 7 of the EEG data (Beta Energy) representing 14% of the data in Group 2 are found to be located outside the specified range, as depicted in Figure 5.61. For Alpha energy in Index 3, 12 of the data, which represents 24% of the data, do no belong to the specificied index as illustrated in Figure 5.62. This scenario might be caused by some variation in the raw EEG data.

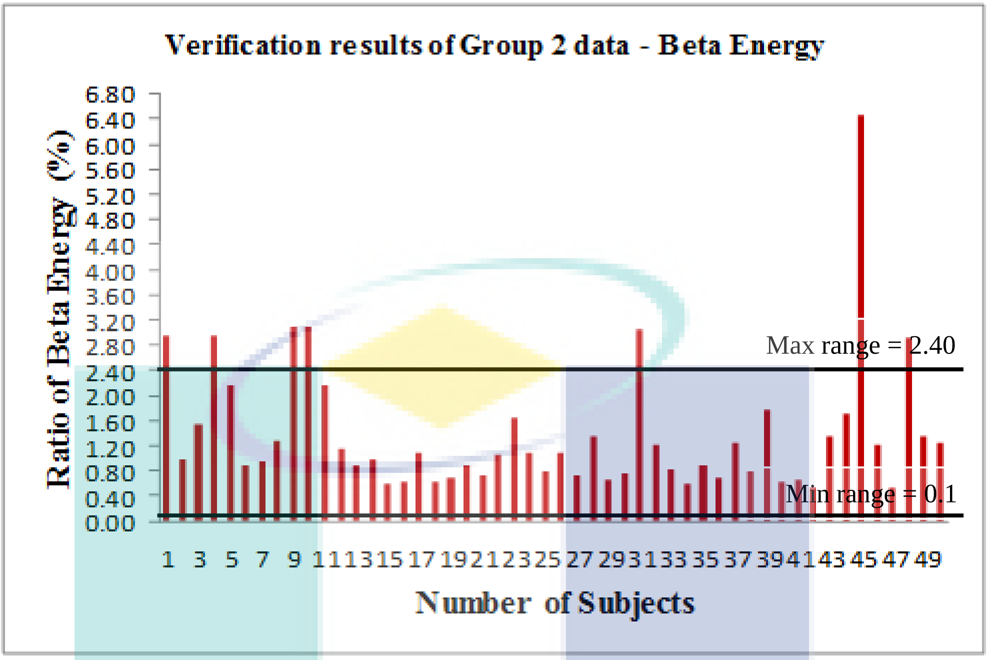


Figure 5.61: Verification Results for Stress Index 3 (Beta Energy) – Group 2

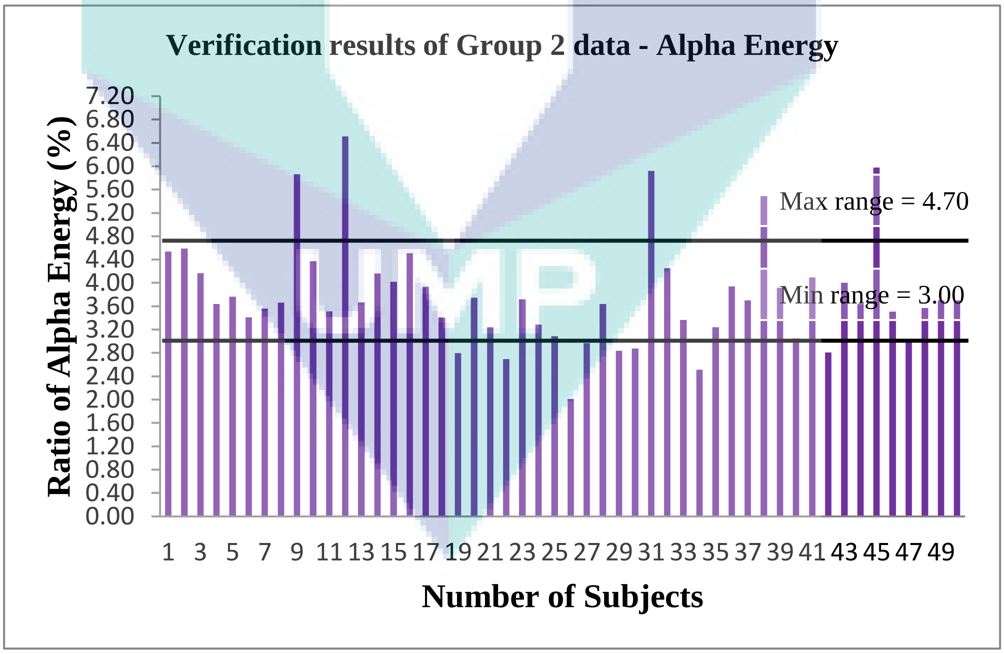


Figure 5.62: Verification Results for Stress Index 3 (Alpha Energy) – Group 2

The verification results for Index 2, which represents Group 3 and Group 4, shows that only 3 EEG data (8%) of the Group 3 do not follow the specified range for Beta energy as shown in Figure 5.63. It is also observed that 7 EEG data of Alpha energy (18%) in Group 3, as shown in Figure 5.64, do not belong to the given stress index. Index 2, which represents moderate stress, is also assigned to participants in Group 4. From the verification process for Group 4, 6 EEG data (15%) are located outside the specified range for Beta energy as shown in Figure 5.65. For Alpha energy in Group 4, 5 EEG data (13%) do not belong to the given stress index as depicted in Figure 5.66.

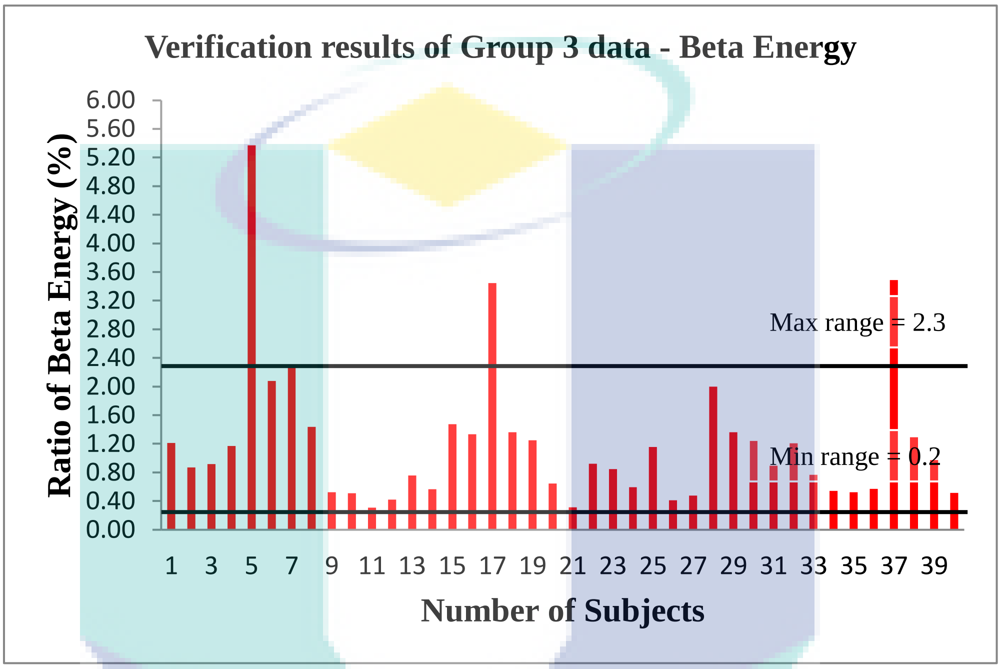


Figure 5.63: Verification Results for Stress Index 2 (Beta Energy) – Group 3

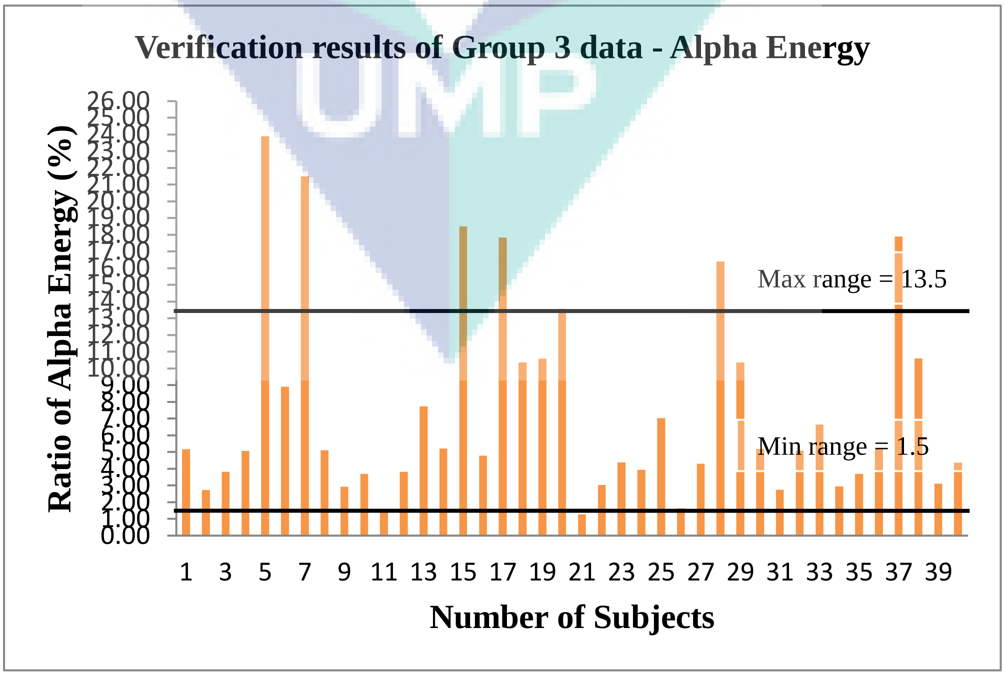


Figure 5.64: Verification Results for Stress Index 2 (Alpha Energy) – Group 3

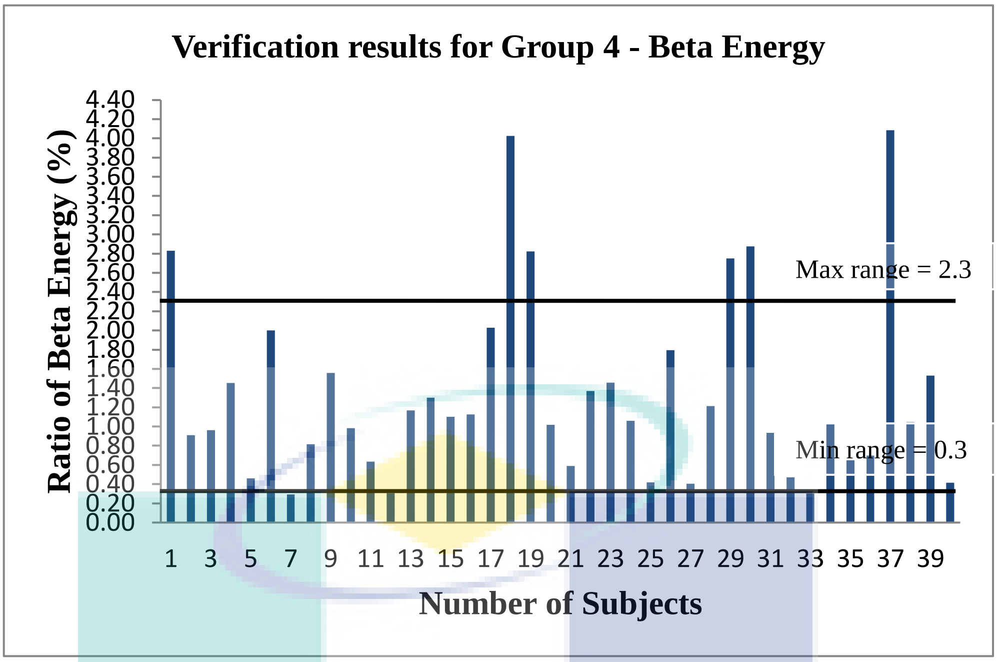


Figure 5.65: Verification Results for Stress Index 2 (Beta Energy) – Group 4

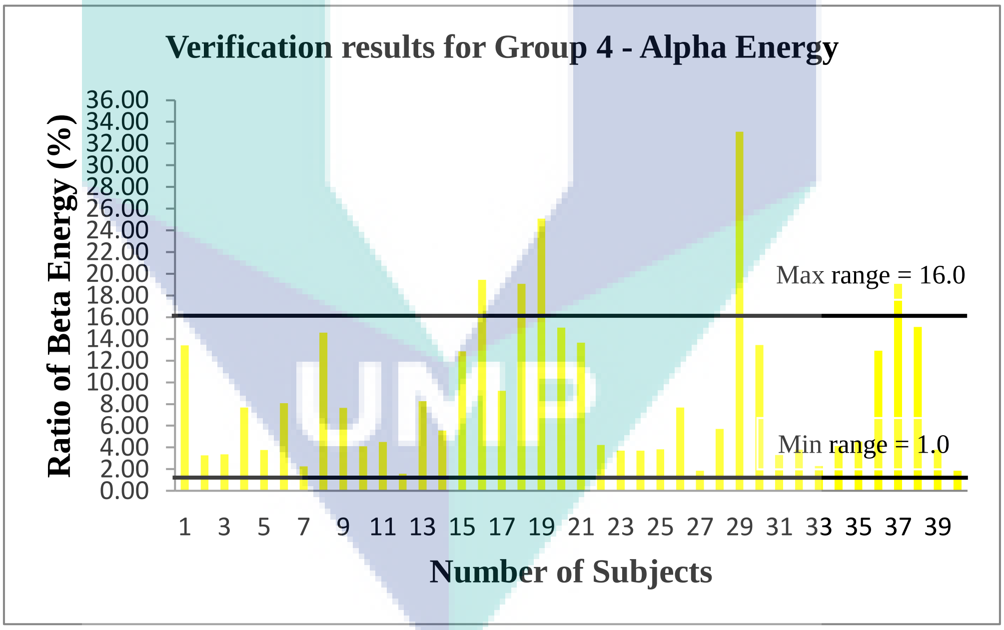


Figure 5.66: Verification Results for Stress Index 2 (Alpha Energy) – Group 4

In order to better understand the results of the verification for the assigned stress index, the number of subjects from each group belonging to the specified index, in terms of Alpha and Beta energy, is counted and plotted as shown by Figure 5.67. Based on the figure, the majority of the EEG data fall into the correct indices with the highest percentage at 92% and the lowest percentage at 80%. For example, in terms of

Beta energy, **86%** of the EEG data in Group 2 is assigned with Index 3, and **80%** of the EEG data of the Alpha energy in Group 2 is assigned with Index 3. Also, stress Index 1 that is assigned to Group 1 is verified to cover 82% of the EEG data of Beta energy belonging to the index, and 92% of the EEG data of Alpha energy belong to the assigned index. Meanwhile, 90% of EEG data of Beta energy for Group 3 belong to the assigned Index 3, and 85% of the data from Group 4 belong to the assigned Index 4. Therefore, based on the verification results, the assigned index can be concluded as a reliable stress index.

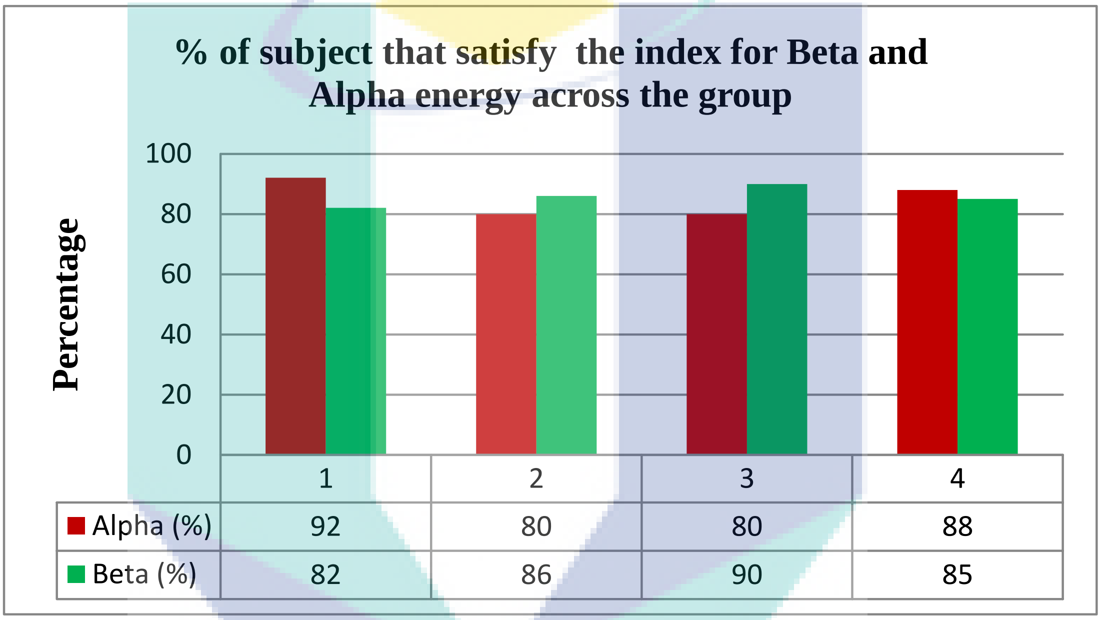


Figure 5.67: Verification Results for Stress Index Across the Groups

Beside the determination of stress index using spectral centroids, the stress index is also established based on the results of the entropy of the EEG data from each participant in the group. Though the range of entropy is small (0.29 – 0.33), it can be used to verify the specified index and the stress verification process used is similar to that used to verify the index based on energy ratio. Figure 5.68 shows 14 out of 50 (28%) EEG data in Group 1 is located outside the assigned stress Index 1, and Figure 5.69 also shows 13 of the EEG data (26%) in Group 2 is located outside the assigned stress Index 3. Meanwhile, 8 out of 40 (20%) EEG data in Group 3 do not belong to the assigned stress Index 2 as illustrated by Figure 5.70, and 33% or 13 out of 40 EEG data in Group 4 do not belong to Index 2 as shown in Figure 5.71. The high

overlapping of data in the verification of stress index might due to the small range of the entropy across the group.

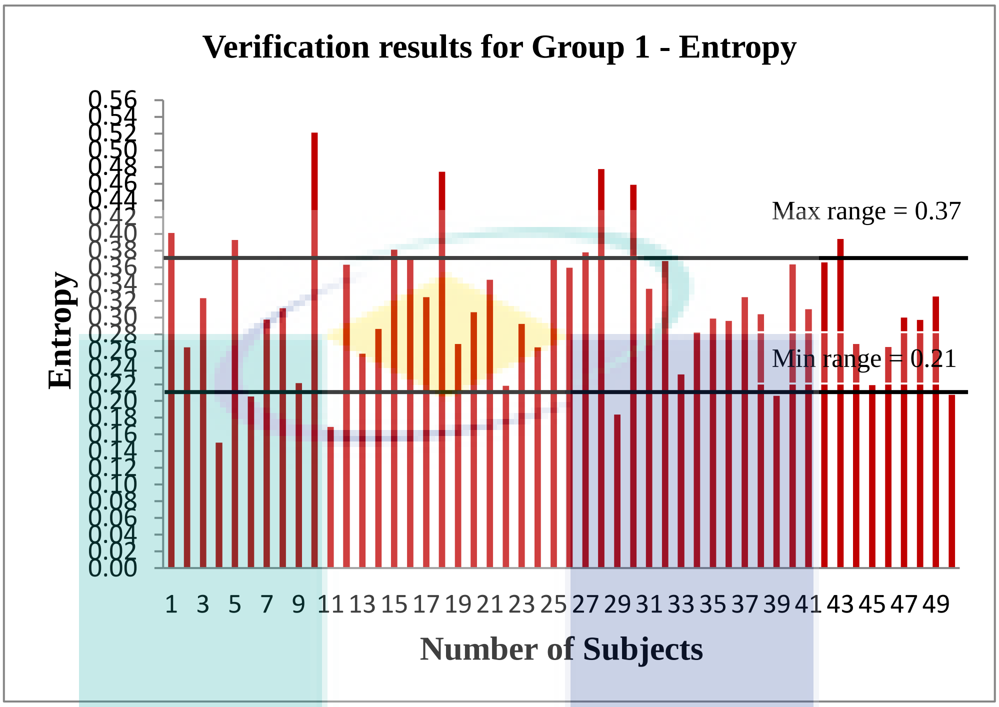


Figure 5.68: Verification Results for Stress Index 1, Group 1 based on Entropy

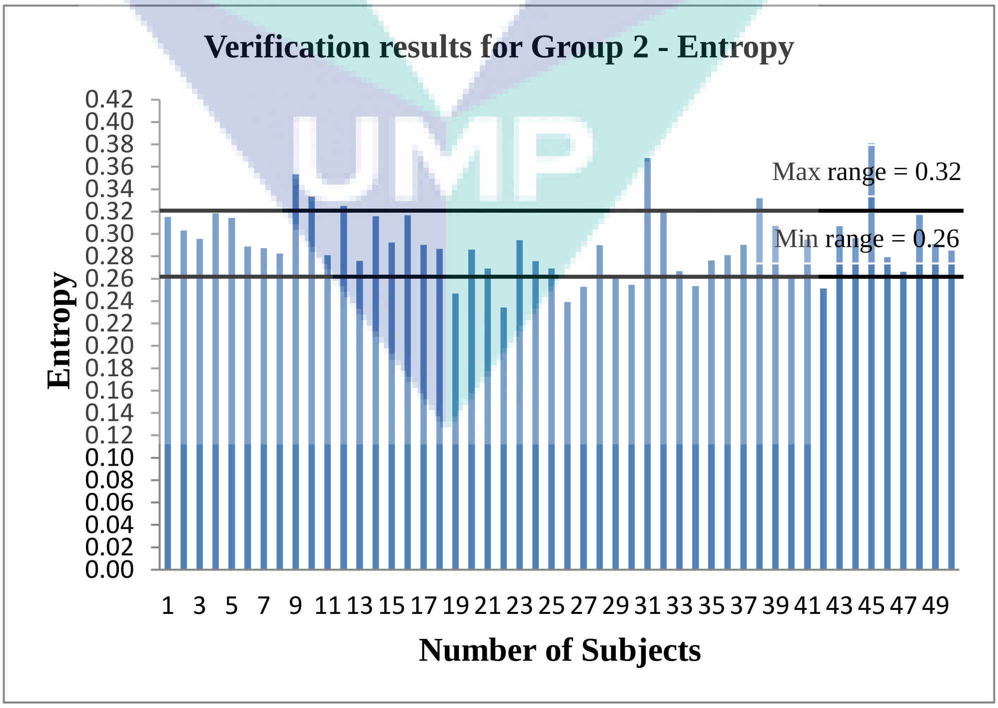


Figure 5.69: Verification Results for Stress Index 3, Group 2 based on Entropy

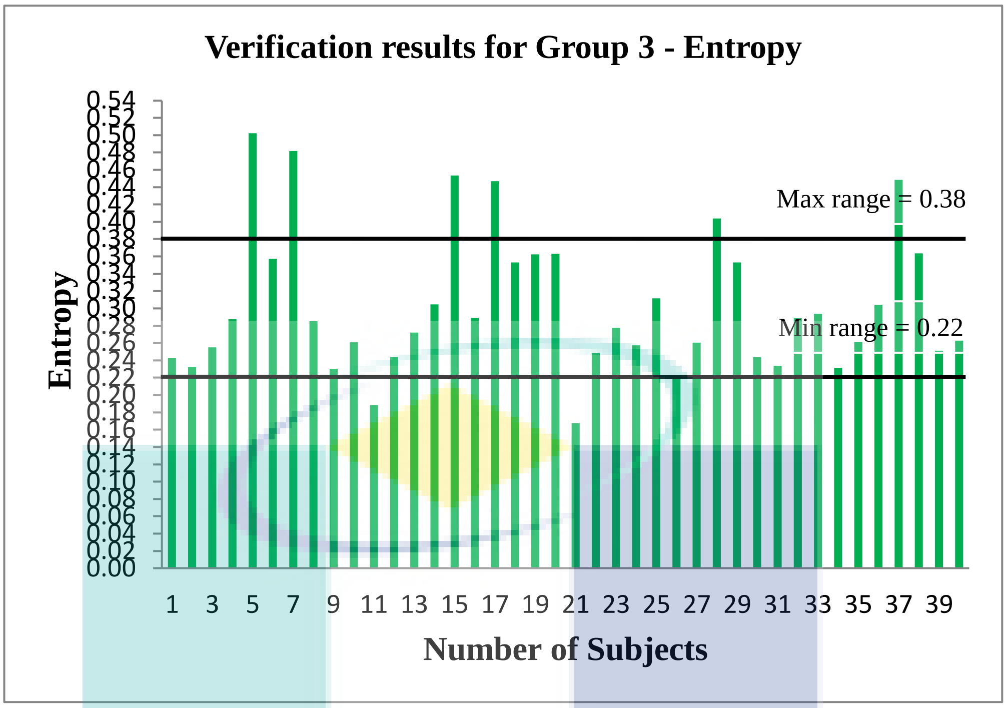


Figure 5.70: Verification Results for Stress Index 2, Group 3 based on Entropy

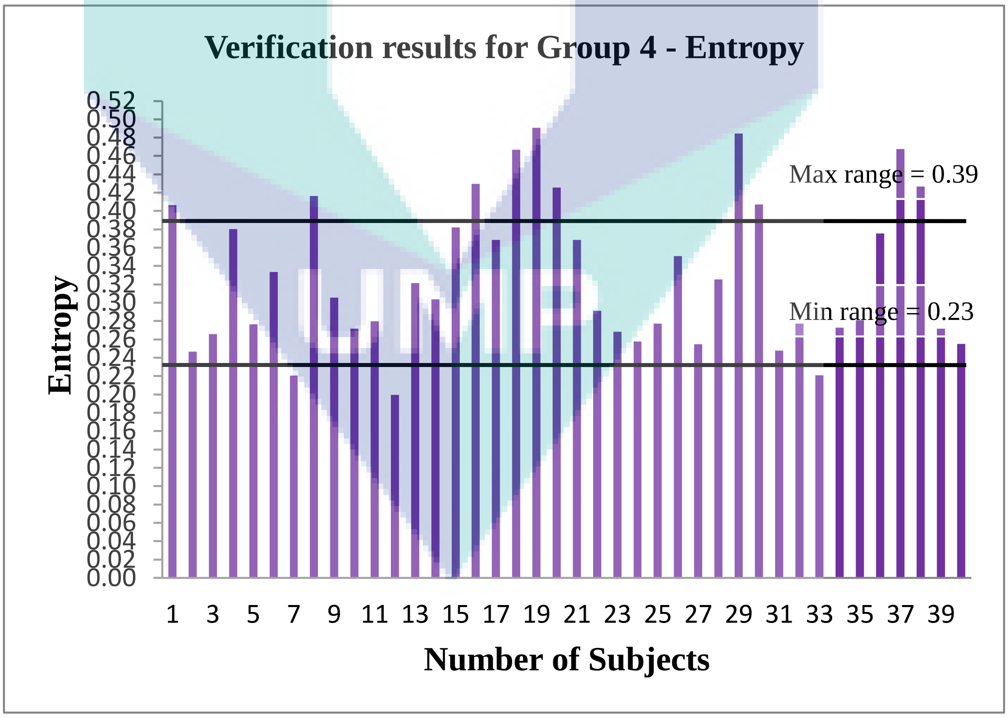


Figure 5.71: Verification Results for Stress Index 2, Group 4 based on Entropy

In order to see the affect of entropy on the verification of the stress index, the percentage of the entropy of the groups that satisfy the proposed stress index is plotted as shown in Figure 5.72. It is observed that Group 3 achieved 80% verification of the assigned stress index, while Group 1 and Group 2 both accomplished 76% and Group 4 attained 70% verification. The lower verification percentage using entropy, compared to that using Alpha and Beta energy, might due to the small range of the entropy across the group, and also the **overlapping** of the EEG data might occur between the groups. The patterns of entropy of the experimental groups indicate that entropy can be used as one of the EEG features to indicate the presence of stress in the groups.

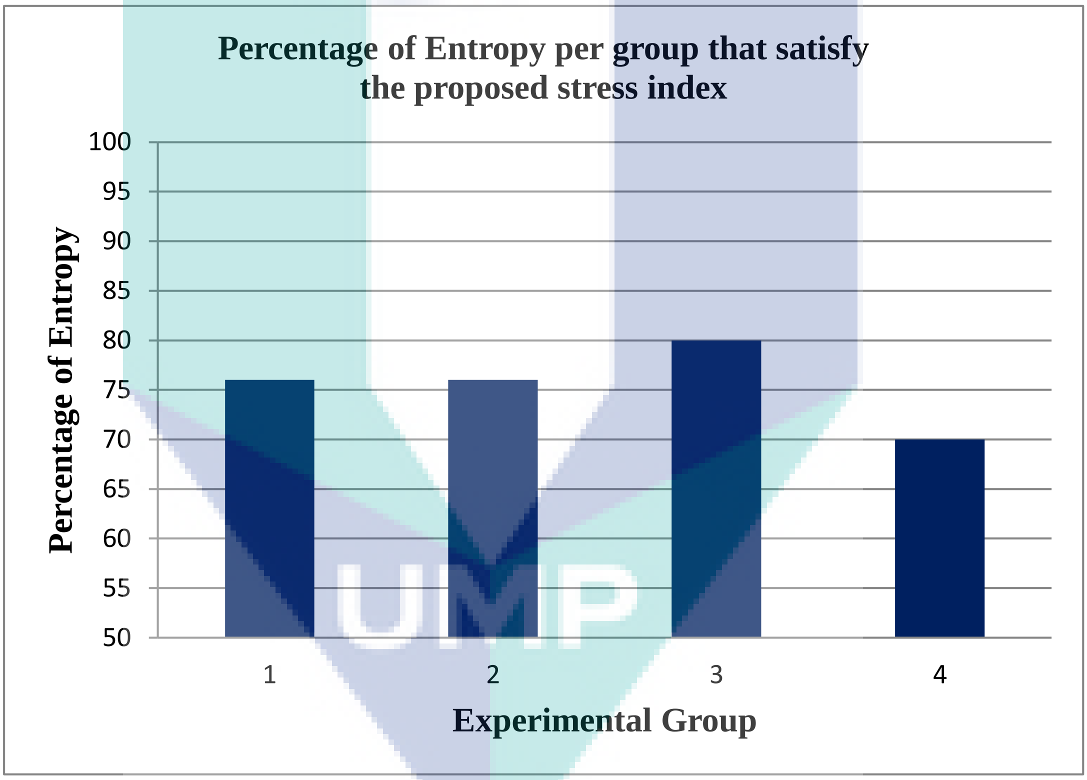


Figure 5.72: Verification Results for Stress Index Across the Groups using Entropy

5.11 CHAPTER SUMMARY

Various experiments were conducted to determine which feature could best detect stress in the group involved in the study. Various features from the spectral analysis of EEG signals are selected as input to k-NN classifier for the classification process. Based on the classification results, the best feature that could detect stress is the relative energy ratio of the spectral density of EEG signals. In addition, the best classification setting is 70:30 training to testing ratio that is used with 180 EEG datasets that were obtained from 140 subjects, while the best classifier configurations to use are Euclidean distance and nearest rule. The performance of the classifier is tested for accuracy, sensitivity and specificity using confusion matrices. To support the k-NN classification results, the EEG features are clustered using FCM and FKM. In addition, ANOVA is applied to verify the robustness of classifier, and then, Pearson correlation is employed to determine the relationship among EEG features in the experimental groups and finally, the classifier is cross-validated using k-fold cross-validation and leave-one-out cross-validation. Since the results of accuracy, sensitivity and specificity of the classifier using the selected EEG features are high including the performance of the cross-validation of the k-NN classifier, the proposed method in determination of stress index can be considered robust. Finally, a stress index is assigned to each group involved in the study based on the centroids, entropy and average energy ratio of Alpha and Beta bands of the EEG signals. From the results of the experiments, Index 1 is assigned to Group 1, Index 2 is assigned to Group 3 and Group 4, and Index 3 is assigned to Group 2.

The index assignment is verified using the Z-score technique and by testing each data from each participant in the groups using Table 5.19 (Z-score specifications) in order to determine how many of them are accurately fall into their respective bins or assigned indices. The results of assigning stress index using Z-score technique as shown in Table 5.20, produce almost the similar index as stated in Table 5.18 (assignment of stress index based on the selected EEG features). For the verification of the proposed index using the range of the energy ratio for each group, majority of the EEG data fall in to their respective indexes with a minimum verification results of **80%**. However, high verification results are not obtained when using entropy which might due to the small range of the entropy which is less than 0.05. The spectral centroid value excluded in the verification process since it is used to act as a target during classification process where the value of the centroid is unique.

CHAPTER SIX
CONCLUSION AND RECOMMENDATIONS

6.1 CONCLUSION

This thesis presents the results of a study on the generation and performance of a stress index based on selected human EEG features using non-parametric technique by applying spectral centroid on the relative energy ratio of EEG Alpha and Beta band. The level of stress of a person is numerically represented as Index 1 for low stress, Index 2 for moderate stress and Index 3 for high stress.

The experimental results of the study confirmed that the proposed EEG features extraction technique and analysis using non-parametric (spectral) technique, capable of revealing the vital features that are buried in the non-stationary EEG signals that can be employed to determine human stress level. The EEG features extracted and analyzed are asymmetry ratio, energy ratio, spectral centroids and entropy, and these features were selected in this study as they can describe the change in neural activities precisely when the brain is stimulated with certain task. Besides, the selected features are widely used in recognizing human emotions and stress has been categorized as negative emotions. However, asymmetry ratio of the signals was not used in generating stress index due to its poor classification results. In addition, only the extracted features from EEG Beta and Alpha bands were analyzed and classified, as previous studies have found that the stress pattern can be indicated by an increase in Beta activity and a decrease in Alpha activity.

This study has succeeded in producing a reliable stress index based on the high accuracy rate, sensitivity rate and specificity rate given by the k-NN classifier with small classification errors. The overall classification accuracy of 88.89%, classification sensitivity of 86.67% and classification specificity of 100% were obtained.

Besides, the classification performance were measured at 78.89% and 83.5% using 10-fold cross-validation and leave-one-out cross-validation techniques, respectively. The good verification results of the stress index using Z-score technique have confirmed the realiability and solidness of the index. The verification of the index was done on the energy ratios of Beta band and Alpha band and on the entropy of all groups. From the verification results, the majority of the EEG data fall into the proposed index with a minimum verification results of 86% for Index 3, 85% to 90% for Index 2 and 82% for Index 1 in terms of energy ratios of Beta band. Meanwhile, using energy ratios of Alpha band, 80% verification results were obtained for Index 3, 80% to 88% for Index 2 and 92% for Index 1. However, the verification results using Shannon's entropy only produces satisfactory results with 76% for Index 1 and Index 3, and 70 to 80% for Index 2.

In order to support the proposal of the stress index assignment, the selected features from the EEG signals were analyzed using ANOVA and Pearson correlation to check the significant difference within the group and between groups and linear correlation of the selected EEG features (Asymmetry and Relative Energy Ratio of EEG Alpha and Beta bands). The results obtained indicate that energy ratios Alpha and Theta bands have significant difference compared to energy ratios of Beta and Delta bands, but no significant difference is observed on asymmetry ratio of all EEG frequency bands. The results of the ANOVA are validated by the results shown by FCM and FKM clustering. The results from Pearson's correlation suggest that the linear correlation exists between selected EEG features across the experimental groups with moderate strength. The strength of Pearson correlation using asymmetry of Alpha and Beta bands is 0.629 with $p < 0.05$. Meanwhile, the correlation strength for the EEG features of energy ratios of Alpha and Beta band is slightly lower which is at 0.576 with $p < 0.05$. Since the statistical results of the selected EEG features are at moderate level, both features can be used for the classification processes. However, the higher classification accuracy rate was obtained using energy ratios compared to the classification results using asymmetry ratios.

Beside statistical analyses, the assigned stress index are verified using Z-score technique and range of the EEG features in terms of energy ratios of Beta band, energy ratios of Alpha band and the range of entropy of the EEG data from each group.

Even though this study has produced stress index with high accuracy, some improvement works need to be implemented to enhance the reliability of the index. The suggestions for future work are as follows.

6.2 RECOMMENDATIONS FOR FUTURE WORK

The thesis has shown that a reliable stress index based on the energy spectral density of the EEG frequency bands can be implemented. This index can be recommended to be applied in various fields as stated below:-

- Health sector

In this field, the index can be employed to measure the stress level a person using their brain signals. This index can be used not only for personal purpose, but for medical treatment as well where physicians, medical practitioners or psychologists can apply it to measure the stress level of their patients before providing a proper consultation and medication. For a real-time usage, a portable stress indicator device can be designed to measure and monitor human stress. The brain signal will be an input to the device, and using this proposed stress index, the stress index which represent the stress level can be displayed in the device. In order to make the subject feels comfortable and to eliminate unnecessary noises that might disturb the stress index, a wireless EEG sensors might be used.

- Educational sector

In educational sector, the proposed stress index might be used to monitor the stress level of students during lectures or examinations. The monitoring results will provide a proper feedback to the lecturers to improve their teaching styles or students potential. The index also can be applied to measure the stress level of the academicians or lecturers during lecture or preparing examination questions.

- Engineering and Research sector

In this field, as alternative method of analysis, the stress index can be visualized in terms of 2-D or 3-D images to produce a better and comprehensive stress index. This study has been implemented based on the analysis of the brain signals in two difference cognitive states: relax (not doing anything) and answering IQ tests. Therefore, this study can be extended by studying various cognitive states and stimulation to the brain such as measuring stress during mental tasks or taking important examination. Besides, various EEG features can be explored other than asymmetry ratio, energy ratio, spectral centroids and entropy that were used in this study. This study applied threshold technique to remove the noise especially EOG signals by specifying the characteristics of the EEG signals to set-up the threshold parameters. In future, the EEG signals, EOG signals and ECG signals might be combined and used simultaneously. Then, Independent Component Analysis (ICA) technique can be employed to differentiate EEG signals from its interference signals in order to produce really clean EEG signals for the next process or unique patterns from the combined signals can be used for analysis.

REFERENCES

[1] S. Handri, S. Nomura and K. Nakamura, "Detecting Stress Based on the Schedule of an Intermittent Mental Workload Using Physiological Sensor," in *Proceedings of the 11th International Conference on Humans and* Computers, 2008, pp. 123-126.

[2] J. A. Healey and R. W. Picard, Detecting Stress During Real-World Driving Tasks Using Physiological Sensors, *IEEE Transactions on Intelligent Transportation Systems*, vol. 6, 1-19, 2004.

[3] S. H. Seo and J. T. Lee, *Stress and EEG*, Convergence and Hybrid Information Technologies, M. Crisan, InTech, 2010, pp. 413-426.

[4] K. S. Rahnuma, A. Wahab, N. Kamaruddin and H. Majid, "EEG Analysis for Understanding Stress Based on Affective Model Basis Function," in *Proceedings of the 15th International Symposium on Consumer Electronics*, 2011, pp. 592-597.

[5] S. Reisman, Measurement of Physiological Stress, in *Proceedings of the IEEE Bioengineering Conference*, 1997, Northeast, USA, pp. 21-23.

[6] M. Trimble and M. S. George, *Biological Psychiatry*. UK:John Wiley & Sons, 2010.

[7] S. Sanei and J. A. Chambers, *EEG Signal Processing*. England: John Wiley & Sons, 2007.

[8] E. R. Kandel, *Principles of Neural* Science. McGraw Hill, 2000.

[9] M. S. Avidan, L. Zhang, B. A. Burnside, K. J. Finkel, A. C. Searleman, J. A. Selvidge, L. Saagar, M. S. Turner, S. Rao, M. Bottros and C. Hantler, "Anesthesia Awareness and Bispectral Index," *The New England Journal of Medicine*, vol. 358, pp. 1097-1108, 2008.

[10] S. Kar, M. Bhagat and A. Routray, "EEG signal analysis for the assessment and quantification on driver's fatigue", *Journal of Transportation Research Part F*, vol. 13, 297-306, 2010.

[11] G. Borghini, L. Astolfi, G. Vecchiato, D. Mattia, and F. Babiloni, "Measuring *neurophysiological* signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness," *Neuroscience and Biobehavioral Reviews*, pp. 1-18, 2012.

[12] K. S. Park, H. Choi, K. J. Lee, J. Y. Lee, K. O. An and E. J. Kim, "Emotion Recognition based on the asymmetric left and right activation," *International Journal of Medicine and Medical Sciences*, vol. 3, pp. 201-209, 2011.

[13] T. Zhuang, H. Zhao and Z. Tang, "A Study of Brainwave Entrainment Based on EEG Brain Dynamics," *Journal of Computer and Information Science*, vol. 2, pp. 80-86, 2009.

[14] I. M. Rosso, A. D. Young, L. A. Femia, and D. A. Yurgelun-Todd, "Cognitive and Emotional Components of Frontal Lobe Functioning in Childhood and Adolescence," *Annals of the New York Academy of Sciences*, vol. 1021, pp. 355-362, 2006.

[15] T. Zhuang, H. Zhao and Z. Tang, "A Study of Brainwave Entrainment Based on EEG Brain Dynamics," *Journal of Computer and Information Science*, vol. 2, pp. 80-86, 2009.

[16] M. Teplan, "Fundamentals of EEG Measurement", *Measurement Science Review*, vol.2, pp. 1-11, 2002.

[17] C. Peter and A. Herbon, "Emotions in Ambient Intelligence – An Experiment on How to Measure Affective States," *Emotion Representation and Physiology Assignments in Digital Systems in Interacting with Computers*, vol. 18, no. 2, pp. 139-170, 2006.

[18] D. Nie, X. W. Wang, L. C. Shi, and B. L. Lu, "EEG-based Emotion Recognition During Watching Movies," in *Proceedings of the 5th IEEE EMBS on Neural Engineering*, 2011, pp. 667-670.

[19] D. O. Bos, "EEG-based Emotion Recognition,"*The influence of Visual and Auditory Stimuli University*, 2006.

[20] A. S. AlMejrad, "Human Emotions Detection using Brain Wave signals," *European Journal of Scientific Research*, vol. 44, no. 4, pp. 640-659, 2010.

[21] J. A. Coan, J. J. B. Allen and P. E. McKnight, "A capability model of differences In frontal EEG asymmetry," *Journal of Biological Psychology*, vol. 72, pp. 198-207, 2006.

[22] J. A. Coan and J. J. B. Allen, "Frontal EEG Asymmetry as moderator and mediator of emotion," *Biological Psychology*, vol. 67, pp. 7-49, 2004.

[23] R. S. Lewis, "The effect of naturalistic stressors on frontal EEG Asymmetry, Stress and Health," *Biological Psychology*, vol. 75, pp. 239-247, 2007.

[24] J. B. Henriques and R. J. Davidson, “Left Frontal Hypoactivation in Depression,” *Journal of Abnormal Psychology*,” vol. 100, pp. 535-545, 1991.

[25] J. B. Henriques and R. J. Davidson, “Brain Electrical Asymmetries during Cognitive Task Performance in Depressed and Nondepressed Subjects,” *Journal of Psychiatry*, vol. 42, pp. 1039-1050, 1997.

[26] D. J. A. Smit, D. Posthuma, D. I. Boomsma, and E. J. C. De Geus, “The relation between frontal asymmetry and the risk for Anxiety and Depression,” *Journal of Biological Psychology*, vol. 74, pp. 26-33, 2007.

[27] A. J. Niemic and B. J. Lithgow, “Alpha-band Characteristics in EEG Spectrum IndicateReliability of Frontal Brain Asymmetry measures in diagnosis Depression,” in *Proceedings of the IEEE Engineering in Medicine and Biology*, 2005.

[28] J. J. B. Allen, J. A. Coan, and M. Nazarian, “Issues and assumptions on the road from raw signals to metrics of frontal EEG Asymmetry in emotion,” *Journal of Psychology*, vol. 67, pp. 183-218, 2004.

[29] I.Papousek and G. Schulter, “Manipulation of frontal brain asymmetry by cognitive tasks”, *Journal of Brain and Cognition*, vol. 54, pp. 43-51, 2004.

[30] S. Debener, A. Beauducel, D. Nessler, B. Brocke, H. Heilemann, and J.Kayser, "Is Resting Anterior EEG Alpha Asymmetry a Trait Marker for Depression?," *Neuropsychobiology*, vol. 41, pp. 31-37, 2000.

[31] M. Murugappan, N. Ramachandran and Y. Sazali, “Classification of human emotion from EEG using discrete wavelet transform,” *Journal of Biomedical Science and Engineering*, vol. 3, pp. 390-396, 2010.

[32] O. Sourina, Q. Wang, Y. Liu, and M. K. Nguyen, “A Real-time Fractal-based Brain State Recognition from EEG and its Applications,” in *Proceedings of the 4th International Conference on Bio-inspired Systems and Signal Processing*, 2011, pp. 82-91.

[33] R. Khosrowabadi, C. Quek, K. A. Kai, W. T. Sau and M. Heijnen, “A Brain-Computer Interface for classifying EEG correlates of chronic mental stress,” in *Proceedings of the International Joint Conference on Neural Network*, 2011, pp. 75-80.

[34] J. Zhai and A. Barreto, “Stress Detection in Computer Users Band on Digital Signal Processing of Noninvasive Physiological Variables”, in *Proceedings of the 28th IEEE EMBS Annual International Conference*, 2006, pp. 1355-1358.

[35] M. K. Abdullah, K. S. Subari, C. L. J. Leo, and N. N. Ahmad, “Analysis of EEG signals for a Practical Biometric System,” *Journal of World Academy of Science Engineering and Technology*, vol. 68, pp. 1123-1127, 2010.

[36] Z. H. Murat, M. N. Taib, S. Lias, R. S. S. Abdul Kadir, N. Sulaiman, and M. Mustafa, “The Conformity between Brainwave Balancing Index (BBI) using EEG and Psychoanalysis Test,” *International Journal of Simulation: Systems, Science and Technology*, vol. 11, no. 2, pp. 86-92, 2010.

[37] R. W. Sperry, “Some Effects of Disconnecting the Cerebral Hemispheres,” *Division of Biology California Institute of Technology*, California, 1981, pp. 1-9.

[38] P. J. Sorgi, *The 7 Systems of Balance: A Natural Prescription* for Healthy Living in Hectic World Health Communications Incorporated, 2002.

[39] A. P. Anokhin, A. C. Heath, and E. Myers, "Genetic and environmental influences on frontal EEG asymmetry: A twin study," *Biological Psychology*, vol. 71, pp. 289-295, 2006.

[40] P. C. Petrantonakis and L. J. Hadjileontiadis, "A Novel Emotion Elicitation Index Using Frontal Brain Asymmetry for Enhanced EEGBased Emotion Recognition," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 15, pp. 737-746, 2011.

[41] M. A. Staal, “Stress, Cognition, and Human Performance: A Literature Review and Conceptual Framework,” Ames Research Center, Moffet Field, California, Technical Report, 2004.

[42] V. Knott, C. Mahoney, S. Kennedy and K. Evans, “EEG Power, frequency, asymmetry and coherence in male depression”, *Journal of Psychiatry Research: Neuroimaging Section*, vol. 106, pp. 123-140, 2001.

[43] E. Hoffman, “Brain Training Against Stress: Theory, Methods and Results from an Outcome Study,” Stress Report, 2005.

[44] P.A.Hancock, P.A.,Desmond, Stress, Workload and Fatigue. New Jersey: Lawrence Erlbaum Associates, Inc., 2001.

[45] M. Chaouachi, I. Jraidi, and C. Frasson, “Modeling Mental Workload using EEG Features for Intelligent Systems,” in *Proceedings of the International Conference on User Modeling, Adaptation, and Personalization*, 2011, pp. 50-61. The cognitive activation theory of stress,” *Psychoneuroendocrinology*, vol. 29, pp. 567-592, 2004.

[46] J. Lyle E. Bourne and R. A. Yaroush, “Stress and Cognition: A Cognitive Psychological Perspective,” University of Colorado, 2003.

[47] B. Roozendaal, B. S. McEwen, and S. Chattarji, “Stress, memory and the amygdala,” *Nature Reviews Neuroscience*, vol. 10, pp. 423-433, 2009.

[48] Harshpinder and P. Aujla, “Physiological and Psychological Stressors among Working and Non-Working Women,” *Journal of Human Ecology*, vol. 20, no. 2, pp. 121-123, 2006.

[49] R. Subhani, L. Xia, and A. S. Malik, “EEG Signals to Measure Mental Stress,” in *Proceedings of the 2nd International Conference on Behavioral, Cognitive and Psychological Sciences*, 2011, pp. 6-10.

[50] M. Hall, D. J. Buyse, P. D. Nowell, E. A. Nofzinger, P. Houck, C. F. Reynolds, and D. J. Kupfer, “Symptoms of Stress and Depression as Correlate of Sleep in Primary Insomnia,” *Psychosomatic Medicine*, vol. 62, pp. 227-230, 2000.

[51] R. K. Sinha, “An approach to estimate EEG Power Spectrum as an Index of Heat Stress using Back propagation Artificial Neural Network,” *Journal of Medical Engineering Physics*, vol. 29, pp. 120-124, 2001.

[52] H. Ursin and H. Eriksen, “The cognitive activation theory of stress,” *Psychoneuroendocrinology*, vol. 29, pp. 567-592, 2004.

[53] S. Rauf, L. Xia and S. Malik, “EEG signals to measure mental stress,” in *Proceedings of 2nd International Conference on Behavioral, Cognitive and Psyhological Sciences*, pp. 6-10, 2011.

[54] V. A. Nejtek, "High and low emotion events influence emotional stress perceptions and are associated with salivary cortisol response changes in a consecutive stress paradigm," *Psychoneuroendocrinology*, vol. 27, 2002, pp. 337-352.

[55] K. S. Park, “Pattern of Electroencephalography (EEG) change against stress through noise and memorization test,” *International Journal of Medicine and Medical Science*, vol. 3, no. 14, pp. 381-389, 2011.

[56] Rajesh Singla, Dr. Balraj Gupta, “Brain Initiated Interaction”, *Journal on Biomedical Science and Engineering*, vol. 1, pp. 170-172, 2008.

[57] C. Babiloni, A. Brancucci, F. Babiloni, P. Capotosto, F. Carducci, F. Cincotti, L. Arendt-Nielsen, A. C. Chen, and P. M. Rossini, “Anticipatory cortical responses during the expectancy of a predictable painful stimulation: A high-

resolution electroencephalography study," *European Journal of Neuroscience*, vol. 18, no. 6, pp. 1692-1700, 2003.

[58] A. Holm, "Developing neurophysiological metrics for the assessment of mental workload and the functional state of the brain*,"* PhD thesis, Aalto University School of Science and Technology, Finland, 2010.

[59] T. Hayashi, Y. Matsumoto, E. Okamoto, R. Ishii, S. Ukai and K. Shinosaki, "Anterior Brain Activity related to Emotional Stress," in *Proceedings of the World Automation Congress*, 2008, pp. 1-6.

[60] M. Teplan, "Audio-Visual Stimulation and Relaxation, Linear and Nonlinear EEG," Ph.D. thesis, Bratislava, Slovak Republic, 2006.

[61] D. Cvetkovic, "Electromagnetic and Audio-Visual Stimulation of the human brain at extremely low frequency," Ph.D. thesis, RMIT University, Australia, 2005.

[62] K. H. Kim, S. W. Wang and S. R. Kim, "Emotion Recognition System using Short-Term Monitoring of Physiological signals," *Journal of Medical and Biological Engineering and Computing*, vol. 42, pp. 419-427, 2010.

[63] S. Stolc, A. Krakovska, and M. Teplan, "Audiovisual Stimulation of Human Brain, Linear and Nonlinear Measures," *Measurement Science Review*, vol. 3, no. 2, pp. 95-98, 2003.

[64] M. Teplan, A. Krakovska, and S. Stolc, "Direct effects of Audio-Visual Stimulation to EEG," *Computer Methods and Programs in Biomedicine*, vol. 102, pp. 17-24, 2011.

[65] S. A. Hosseini and M. B. Naghibi-Sistani, "Emotion recognition method using entropy analysis of EEG signals," *International Journal of Image Graphics and Signal Processing*, vol. 5, pp. 303-306, 2011.

[66] K. Schaaff and T. Schultz, "Towards and EEG-based Emotion Recognizer for Humanoid Robots," in *Proceedings of the international symposium on Robot and Human Interactive Communication*, 2009, pp. 792-796.

[67] C. Zhao, C. Zheng, M. Zhao and J. Liu, "Physiological Assessment of Driving Mental Fatigue Using Wavelet Packet Energy and Random Forests," *American Journal of Biomedical Sciences* vol., 2, pp. 262-274, 2010.

[68] P. V. D. Haak, R. V. Lon, J. V. D. Meer and L. Rothkrantz, "Stress Assessment of Car-Drivers using EEG-analysis", *International Conference on Computer Systems and Technologies*, 2010, pp. 473-477.

[69] S. Koelstra, M. Pantic, and I. Patras, "DEAP: A Database for Emotion Analysis using Physiological Signals," *IEEE Transactions on Affective Computing*, vol. 3, no. 1, pp. 18-31, 2012.

[70] T. Musha, Y. Terasaki, H. Haque and G. Ivanitsky, "Feature extraction from EEGs Associated with Emotions," *Artifical Life and Robotics*, vol. 1, no. 1, pp. 15-19, 1997.

[71] A. Nozawa and K. Karita, "A Emotion Spectrum Analysis for a daily repetitive mental workload," *Journal of Artificial Life and Robotics*, vol. 15, pp. 316-319, 2010.

[72] T. Hayashi, E. Okamoto, H. Nishimura, H. Inada, R. Ishii, S. Ukai, K. Shinosaki, and Y. Mizuno-Matsumoto, "Beta Activities in EEG Associated with Emotional Stress," *In Medical Sciences and Image Processing(IC-MED)*, vol. 3, no. 1, pp. 57-68, 2009.

[73] J. Raven, "The Raven's Progressive Matrices: Change and Stability over Culture and Time", *Cognitive Psychology*, vol. 41, pp. 1-48, 2000.

[74] A. Akhbas, "Evaluation on the physiological data indicating the dynamic stress level of drivers," *Journal of Scientific Research and Essays*, vol. 6, no. 2, pp. 430-439, 2011.

[75] M. Poulos, M. Rangoussi, N. Alexandris and A. Evangelou, "On the use of EEG features towards person identification via neural networks," in *Proceedings of the International Conference on Acoustic, Speech and Signal Processing*, 1999, pp. 1-35.

[76] S.Ito, Y. Mitsukura, K. Sato, S. Fujisawa and M. Fukumi, "Study on Relationship between Personality and Individual Characteristic of EEG for Personalized BCI," in *Proceedings of the IEEE International Conference on Computational Technology in Electrical and Electronics Engineering (SIBIRCON)*, 2010, pp. 106-111.

[77] T. Lan, A. Adami, D. Erdogmus and M. Pavel, "Estimating Cognitive State using EEG signals," *Journal of Machine Learning*, vol. 4, pp. 1261-1269, 2003.

[78] H. Hinrikus, A. Suhhova, M. Bachmann, K. Aadamsoo, V. Vohma, J. Loss and V. Tuulik, Electroencephalographic Spectral Asymmetry Index for

Detection of Depression, *Journal of Medical Biology Engineering Computing* 47, 1291-1299, 2009.

[79] G. C. Blackhart, J. A.Minnix, and J. P. Kline, “Can EEG Asymmetry patterns predicts future development of anxiety dan depression: A preliminary study,” *Journal of Biological Psychology*, vol. 72, pp. 46-50, 2006.

[80] S. A. Reid, L. M. Duke, and J. J. B. Allen, “Resting Frontal Electroencephalographic Asymmetry in Depression: Inconsistencies suggest the need to identify mediating factors,” *Journal of Psychophysiology*, vol. 35, pp. 389-404, 1998.

[81] Y. Tran, R. A. Thuraisingham, N. Wijesuriya, H. T. Nguyen and A. Craig, Detecting Neural Change during Stress and Fatigue effectively: a comparison of Spectral Analysis and Sample Entropy, in *Proceedings of the IEEE EMBS Conference on Neural Engineering*, 2007, pp. 350-353.

[82] P. Manoilov, “EEG Power Spectrum Analysis during Mental Task Performance,” *International Conference on Computer Systems and Technologies*, 2006, pp. 41-45.

[83] J. R. G. Cooper, J. Al-Muhtadi, R. Ashford, “Smart Spaces with Real-time Physiological Measurement and Mitigation of Sress,” *Pervasive Computing and Applications*, vol. 1, 2008.

[84] Y. S. Kim, H. J. Baek, J. S. Kim, H. S. Lee, J. M. Choi, and K. S. Park, “Helmet based Physiological signal monitoring system, “ *European Journal of Applied* Physiology, vol. 105, no. 3, 2008.

[85] C. Zhao, C. Zheng, M. Zhao and J. Liu, Physiological Assessment of Driving Mental Fatigue Using Wavelet Packet Energy and Random Forests, *American Journal of Biomedical Sciences*, vol. 2, pp. 262-274, 2010.

[86] P. D. Tyson, “Task Related Stress and EEG Alpha Biofeedback,” *Biofeedback and Self-Regulation*, vol. 12, no. 2, pp. 105-118, 1987.

[87] J. F. Cavanagh and J. J. B. Allen, “Multiple aspects of the stress response under social evaluative threat: An electrophysiological investigation,” *Psychoneuroendocrinology*, vol. 33, pp. 41-53, 2008.

[88] D. C. Hammond, “Neurofeedback Treatment for Depression and Anxiety,” *Journal of Adult Development*, vol. 12, no. 2, 2005.

[89] M. Senthimurugan, M. Latha and N. Malmurugan, "Classification in EEG-based Brain Computer Interfaces Using Inverse Model", *International Journal of Computer Theory and Engineering,* vol. 3, no. 2, pp. 274-276, 2011.

[90] B. Hu, D. Majoe, M. Ratcliffe, Y. Qi, Q. Zhao, H. Peng, D. Fan, F. Zheng, M. Jackson, and P. Moore, "EEG-based Cognitive Interfaes for Ubiquitos Application: Development and Challenge," *Journal of Intelligent Systems,* 2011.

[91] S. Cohen, T. Kamarck, and R. Mermelstein, "A Global Measure of Perceived Stress," *Journal of Health and Social Behavior*, vol. 24, pp. 385-396, 1983.

[92] H S. Cohen, D. A. Tyrell, and A. P, "Negative life events, perceived stress, negative affect and susceptibility to the common cold," *Journal of Personality and Social Psychology*, pp. 131-140, 1993.

[93] S. Cohen, T. Kamarck and R. Mermelstein, "A Global Measure of Perceived Stress", *Journal of Health and Social Behavior*, vol. 24, pp. 385-396, 1983.

[94] M. Murugappan, M. Rizon, P. Nagarajan, and S. Yaacob, "Inferring of Human Emotional States using Multichannel EEG," *European Journal of Scientific Research*, vol. 48, no. 2., pp. 281-299, 2010.

[95] J. N. Kannathal, M. L. Choo, U. R. Acharya and P. K. Sadasivan, "Entropies for detection of epilepsy in EEG," *International Journal of Computer Methods and Programs in Biomedicine*, vol. 80, pp. 187-194, 2005.

[96] J. B. Nitschke, W. Heller, M. A. Etienne, and G. A. Miller, "Prefrontal cortex activity differentiates processes affecting memory in depression, "*Biological Psychology*, vol. 67, pp. 125-143, 2004.

[97] P. J. Deldin and, P. Chiu, "Cognitive restructuring and EEG in major depression," *Biological Psychology*, vol. 70, pp. 141-151, 2005.

[98] A. Subasi and M. I. Gursoy, "EEG signal classification using PCA, ICA, LDA, and support vector machines," *Expert Systems with Applications*, vol. 37, no. 12, pp. 8659-8666, 2010.

[99] A. Bashashati, M. Fatourechi and R. K. Ward, "A survey of signal processing algorithms in brain-computer interfaces based on electrical brain signals," *Journal of Neural Engineering*, vol. 4, pp. 32-57, 2007.

[100] V. Sakkalis, T. Cassar, M. Zervakis, K. P. Camilleri, S. G. Fabri, C. Bigan, E. Karakonstantakl, and S. Micheloyannis, "Parametic and nonparametric EEG analysis for the evaluation of EEG activity in young children with controlled

Epilepsy," *Journal of Computational Intelligence and Neuroscience*, pp. 1-15, 2008.

[101] M. Murugappan, R. Nagarajan and S. Yaacob, "Modified Energy Base Time-Frequency Features for Classifying Human Emotions using EEG," in *Proceedings of the International Conference on Man-Machine Systems*, 2009, pp. 1-5.

[102] T. P. Jung, S. Makeig, M. Stensmo, and T. J. Sejnowski, "Estimating Alertness from EEG Power Spectrum," IEEE Transaction in Biomedical Engineering, vol. 44, no. 1, pp. 60-69, 1997.

[103] Z. Iscan, Z. Dokur and T. Demiralp, "Classification of Electroencephalogram signals with combined time and frequencies features," *Expert Systems with Applications*, vol. 38, pp. 10499-10505, 2011.

[104] P. Rani, J. Sims, R. Brackin and N. Sarkar, "Online Stress Detection Using Psychophysiological Signal for Implicit Human-Robot Cooperation," *International Journal of Robotica*, vol. 20, no. 6, pp. 673-686, 2002.

[105] Shuren Qin, Zhong Ji, "Multi-Resolution Time-Frequency Analysis foor Detection of Rhythms of EEG Signals," in *Proceedings of the IEEE 11th Digital Signal Processing Workshop*, pp. 338-341, 2004.

[106] X. Zhang, L. Yin, and W. Wang, "Wavelet time-frequency analysis of Electro-encephalogram (EEG) Processing," *International Journal of Advanced Computer Science and Application*, vol. 1, no. 5, pp. 1-5, 2010.

[107] B. G. Xu and A. G. Song, "Pattern recognition of motor imagery EEG using wavelet transform, *Journal of Biomedical Science and Engineering*, vol. 1, pp. 64-67, 2008.

[108] I. Omerhodzic, S. Avdakovic, A. Nuhanovic, and K. Dizdarevic, "Energy Distibution of EEG signals: EEG Signal Wavelet-Neural Network Classifier," *International Journal of Biological and Life Sciences*, vol. 6, no. 4, pp. 212-215, 2010.

[109] J. Rafiee, M. A. Rafiee, N. Prause and M. P. Schoen, "Wavelet basis functions in Biomedical signal processing", *Expert Systems*, vol. 38, pp. 6190-6201.

[110] K. S. Rahnuma, A. Wahab, N. Kamaruddin and H. Majid, "EEG Analysis for Understanding Stress Based on Affective Model Basis function, in *Proceedings of 1EEE 15th International Symposium on Consumer Electronics*, 2011, pp. 592-597.

[111] S. N. Sarbadhikari and K. Chakrabarty, “Chaos in the brain: a short review alluding the epilepsy, depression exercise and lateralization,” *Medical Engineering & Physics*, vol. 23, pp. 445-455, 2001.

[112] A. B. Ritter, S. Reisman, and B. B. Michniak, *Biomedical Engineering Principles*. New Jersey: CRC Press, 2005.

[113] Farida, Suprijanto, Robinsar, Gunawan, “Time-Frequency analysis of EEG signal response due to simple Acupuncture Stimulation,” in *Proceedings of World Academy of Science, Engineering and Technology*, vol. 38, pp. 317-323, 2009.

[114] K. Q. Shen, S. J. Ong, X. P. Li, Z. Hui, and E. P. V. Wilder-Smith, “A Feature Selection Method for Multilevel Mental Fatigue EEG Classification,” *IEEE Transaction of Biomedical Engineering*, vol. 54, no. 7, 2007.

[115] W. Abdulla and L. Wong, “Neonatal EEG signal characteristics using time frequency analysis,” *Physic A: Statistical Mechanics and its Applications*, vol. 390, no. 6, pp. 1096-1110, 2011.

[116] C. Zhao, M. Zhao, J. Liu, and C. Zheng, “Electroencephalogram and electrocardiograph assessment of mental fatigue in driving simulator,” *Journal of Accident Analysis Prevention*, vol. 45, pp. 83-90, 2012.

[117] Y. Tran, R. A. Thuraisingham, N. Wijesuriya, H. T. Nguyen and A. Craig, “Detecting Neural Change during Stress and Fatigue effectively: a comparison of Spectral Analysis and Sample Entropy,” in *Proceedings of the IEEE EMBS Conference on Neural* Engineering, 2007, pp. 350-353.

[118] M. Ermes, “Methods for the Classification of Biosignals Applied to the Detection of Epileptic Waveforms and to the Recognition of Physical Activity,” Ph.D. thesis, Tampere University of Technology, Tampere, Finland, 2009.

[119] S. Kurimori and T. Kakizaki, “Evaluation of Work Stress Using Psychological and Physiological Measures of Mental Activity in a Paced Calculating Task,” *Industrial Health*, vol. 33, pp. 7-22, 1995.

[120] O. A. Rosso, M. T. Martin, and A. Plastino, “Brain electrical activity analysis using wavelet-based information tools,” *Physic A*, vol. 313, pp. 587, 2002.

[121] P. He, G. Wilson, and C. Russell, “Removal of ocular artifacts from electro-encephalogram by adaptive filtering,” *Medical Biology Engineering Computation*, vol. 42, pp. 407-412, 2004.

[122] G. L. Wallstrom, R. E. Kassb, A. Millerc, J. F. Cohnd, and N. A. Foxe, "Automatic Correction of Ocular Artifacts in the EEG: A comparison of regression-based and component-based methods," *International Journal of Psychophysiology*, vol. 53, pp. 105-119, 2004.

[123] E. Kroupi, A. Yazdani, J. M. Vesin, and T. Ebrahimi, "Ocular Artifact Removal from EEG: A Comparison of Subspace Projection and Adaptive Filtering Methods," in *Proceedings of the 19th European Signal Proceedings Conference*, pp. 1-5, 2011.

[124] C. Guerrero-Mosquera and A. N. Vazquez, "Automatic Removal of Ocular Artifacts from EEG Data using Adaptive Filtering and Independent Component Analysis," in *Proceedings of the 17th European Signal Processing Conference*, 2009, pp. 2317-2321.

[125] Md. K. I. Molla, T. Tanaka, T. M. Rutkowski, and A. Cichocki, "Separation of EOG Artifacts from EEG Signals using Bivariate EMD (BEMD()," in *Proceedings of international conference in Acoustics, Speech and Signal Processing*, 2010, pp. 562-565.

[126] V. Krishnaveni, S. Jayaraman, L. Anitha, and K. Ramadoss, "Removal of Ocular Artifacts from EEG using adaptive thresholding of wavelet coefficients," *Journal of Neural Engineering*, vol. 3, pp. 338-346, 2006.

[127] J. L. Kenemans, P. C. Molenaar, M. N. Verbaten and J. L. Slangen, "Removal of the ocular artifact from EEG: A comparison of time and frequency domain methods with simulated and real data," *Psychophysiology*, vol. 28, pp. 114-121, 1991.

[128] S. Sun and C. Zhang, "Assessing Features for Electroencephalographic Signal Categorization," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, 2005, pp. 417-420.

[129] C. Miyamoto, S. Baba and I. Nakanishi, "Biometric Person Authentication Using New Spectral Features of Electroencephalogram (EEG)," in *Proceedings of the International Symposium on Intelligent Signal Processing and Communication Systems*, 2008, pp. 1-4.

[130] K. K. Paliwal, "Spectral Subband Centroids Features for Speech Recognition," in *Proceedings of the International Conference on Acoustics, Speech and Signal Processing*, 1998, pp. 617-620.

[131] J. S. Seo, M. Jin, S. Lee, D. Jang, S. Lee and C. D. Yoo, "Audio Fingerprinting

Based on Normalized Spectral Subband Centroids," in *Proceedings of the International Conference on Acoustics, Speech and Signal* Processing, 2005, pp. 213-216.

[132] M. L. Massar, M. Fickus, E. Bryan, D. T. Petkie and A. J. Terzuoli Jr., Fast Computation of Spectral Centroids, *Journal of Advanced Computing Mathematics*, 2010.

[133] B. Gajic and K. K. Paliwal, "Robust Parameters for Speech Recognition Based on Subband Spectral Centroid Histogram," in *Proceedings of European Speech*, 2001.

[134] J. Chen, Y. Huang, Q. Li, and K. K. Paliwal, "Recognition of Noisy Speech Using Dynamic Spectral Subband Centroids." *IEEE Signal Processing Letters*, vol. 11, no. 2, pp. 258-261, 2004.

[135] A. Holm, K. Lukander, J. Korpela and M. Sallinen, "Estimating Brain Load from the EEG", *The Scientific World Journal*, vol. 9, pp. 639-651, 2009.

[136] A. Gevin and M. E. Smith, "Neurophysiological measures of cognitive workload during human-computer interaction," *Theoretical Issues in Ergonomics Science*, vol. 4, no. 1-2, pp. 113-131, 2003.

[137] H. C. Shih and S. H. Fang, "Assessment of Psychological Stress using Heart Rate Interval and Plethysmography Wave Amplitude," Yuan Ze Univesity, Taiwan, Project Report, 2008.

[138] G. F. Wilson and C. A. Russell, "Real-time Assessment of Mental Workload using Psychophysiological Measures and Artificial Neural Networks," *Human Factor: The Journal of the Human Factors and Ergonomics Society*, vol. 45, pp. 635-643, 2003.

[139] A. Nassef, M. Mahfouf, D. A. Linkens, E. Elsamahy, A. Roberts, P. Nickel, G. R. J. Hockey and G. Panoutsos, "The Assessment of HRV and Task Load Index (TLI) as Physiological Markers for Physical Stress," *Internal Federation of Medical and Biological Engineering*, 2009, pp. 146-149.

[140] A. Picot, S. Charbonnier and A. Caplier, "On-line Automatic Detection of Driver Drowsiness using a single Electroencephalographic Channel", in *Proceedings of the International Conference of IEEE Engineering in Medicine and Biology Society (EMBC)*, Vancouver, Canada, 2008.

[141] P. Antonenko, F. Paas, R. Grabner, and T. van Gog, "Using Electroencephalography to Measure Cognitive Load," *International Journal of Education Psychological*, vo. 22, pp. 425-438, 2010.

[142] I. Jraidi, A. Heraz, M. Chaouachi and C. Frasson, "New Architecture of a Multi Agent System which Measures the Learner Brainwaves to Predict his Stress Level Variation," in *Proceedings of the World Conference on E-Learning in Corporate, Government, Healthcare and Higher Education*, 2009, pp. 2726-2733.

[143] T. Yamakoshi, P. Rolfe, Y. Yamakoshi and H. Hirose, "A novel physiological index for Driver's Activation State (DAS)," *Journal of Transportation Research Part C*, vol. 17, pp. 69-80, 2000.

[144] C. Berka, D. J. Levendowski, M. M. Cvetinovic, M. M. Petrovic, G. Davis, M. N. Lumicao, V. T. Zivkovic, M. V. Popovic, and R. Olmstead, "Real-time Analysis of EEG Indexes of Alertness, Cognition, Memory acquired with a wireless EEG headset," *International Journal of Human-Computer Interaction*, vol. 17, no. 2, pp. 151-170, 2004.

[145] T. P. Exarchos, A. T. Tzallas, D. I. Fotiadis, S. Konitsiosis, and S. Gianopoulos, "EEG Transient Event Detection and Classification using Association Rules," *IEEE Transactions on Information Technology in Biomedicine*, vol. 10, no. 3, pp. 451-457, 2006.

[146] A. Mahdi, A. Razali and A. AlWakil, "Comparison of Fuzzy Diagnosis with k-Nearest Neighbor and Naïve Bayes Classifiers in Disease Diagnosis," *International Journal of Artificial Intelligence and Neuroscience*, vol. 2, pp. 58-66, 2011.

[147] P. Abhang, S. Rao, B. W. Gawali, and P. Rokade, "Emotion Recognition using Speech and EEG signal – A Review," *International Journal of Computer Applications*, vol. 15, no. 3, pp. 37-40, 2011.

[148] K.P. Mohandas, M. Gerropati, "Artificial Neural Networks for Classification of EEG signals for Brain Computer Interface," in *Proceedings of the FAE Symposium*, 2003, pp. 1-6.

[149] K. Nakayama, Y. Kaneda and A. Hirano, "A Brain Computer Interface Based on FFT and Multilayer Neural Network – Feature Extraction and Generalization," in *Proceedings of the International Symposium on Intelligent Signal Processing and Communication Systems*, 2007, pp. 826-829.

[150] T. Y. Chai, S.S. Woo, M. Rizon and CS Tan, “Classification of Human Emotion from EEG signals using Statistical Features and Neural Network,” *International Journal of Integrated Engineering (Issue on Electrical Engineering)*, vol. 1, no. 3, pp. 71-79, 2010.

[151] M. D. Duta, “The study of vigilance using Neural Network analysis,” PhD thesis, Oxford University, 1998.

[152] C. Castellaro, G. Favaro, A. Casabgrande, D. V. Puthenparampil, and C. F. Salimbeni, “An Artificial Intelligence Approach to Classify and Analyze EEG Traces,” *International Journal of Clinical Neurophysiology*, vol. 32, pp. 1-22, 2002.

[153] C. J. Lin and M.H. Hsieh, “Classification of mental task from EEG data using Neural Networks based on Particle Swarm Optimization,” *Neurocomputing*, vol. 72, pp. 1121-1130, 2009.

[154] A. Subasi, M. Kiymik, A. Alkan, and E. Koklukaya, “Neural Network Classification of EEG Signals by using LR with MLE Preprocessing for Epileptic Seizure Detection,” *Mathematical and Computational Applications*, vol. 10, no. 1, pp. 57-70, 2005.

[155] A. Subasi, “Automatic recognition of Alertness level from EEG by using Neural Network and Wavelet coefficient,” *Expert Systems with Applications*, vol. 28, 2005.

[156] S. Chandaka, A. Chatterjee, and S. Munshi, “Cross-correlation aided support vector machine classifier for classification of EEG signals,” *Expert Systems with Applications*, vol. 36, no. 2, pp. 1329-1336, 2009.

[157] L. Yuan-Pin, W. Chi-Hong, W. Tien-Lin, J. Shyh-Kang, and C. Jyh-Horng, "Support vector machine for EEG signal classification during listening to emotional music," in *Proceedings of the Multimedia Signal Processing*, 2008, pp. 127-130.

[158] P. Pujari and J. B. Gupta, “Improving Classification Accuracy by Using Feature Selection and Ensemble Model,” *International Journal of Soft Computing and Engineering (IJSCE)*, vol. 2, no. 2, pp. 380-386, 2012.

[159] M. Murugappan, M. Rizon, R. Nagarajan, and S. Yaacob, “FCM Clustering of Human Emotions using Wavelet based features from EEG,”, *Biomedical Soft Computing and Human Sciences*, vol. 14, no. 2, pp. 35-40, 2009.

[160] M. Murugappan, M. Rizon, R. Nagarajan, S. Yaacob, I. Zunaidi, and D. Hazry, "ÉEG feature extraction for classifying emotions using FCM and FKM," *International Journal of Computers and Communications*, vol. 1, no.2, pp. 21-25, 2007.

[161] M. Ashtiyani, S. Asadi and P. M. Birgani, ICA-Based EEG Classification Using Fuzzy C-Mean Algorithm, in *Proceedings of the 3rd International Conference on Telecommunication Technology and Applications*, 2008.

[162] M. Sabeti, M. H. Sadreddini, and G. W. Price, "Fuzzy Accuracy-based Classifier Systems for EEG classification of Schizophrenic," *First Join Congress on Fuzzy and Intelligent Systems*, 2007.

[163] S. Chiu, "Fuzzy model identification based on cluster estimation, "*Journal of intelligent and Fuzzy systems*, vol. 2, pp. 267-278, 1994.

[164] A. Alkan and S. B. Akben, "Use of K-Means clustering in migraine detection by using EEG records under flash stimulation," *International Journal of the Physical Sciences*, vol. 6, pp. 641-650, 2011.

[165] M. Jiang and Z. Wang, "A Method for Stress Detection Based on FCM Algorithm," *Proceedings of the 2nd International Congress on Image and Signal Processing*, 2009, pp. 1-5.

[166] A. de Santos, C. S. Avilla, J. G. Casanova, and G. B. del Pozo, "A stress detection systems based on physiological signals and fuzzy logic, IEEE Trans, *on Industrial Electronis*, vol. 58, pp. 4857-4865, 2011.

[167] J. Murakami, S. I. Ito, Y. Mitsukura, C. Jianting, and M. Fukumi, "A design of the EEG features detection and condition classification," in *Proceedings of the SICE Annual Conference*, 2007, pp. 2798-2803.

[168] A. Mahdi, A. Razali and A. AlWakil, "Comparison of Fuzzy Diagnosis with k-Nearest Neighbor and Naïve Bayes Classifiers in Disease Diagnosis," *International Journal of Artificial Intelligence and Neuroscience*, vol. 2, pp. 58-66, 2011.

[169] S. Visa, B. Ramsay, A. Ralescu, and E.van der Knaap, "Confusion Matrix-based Feature Selection," in *Proceedings of the 22th International Conference on Artificial Intelligence and Cognitive Science*, 2011, pp. 1-8.

[170] M. Mullin and R. Sukthankar, "Complete Cross-Validation for Nearest Neighbor Classifiers," in *Proceedings of the 17th International Conference on Machine Learning*, 2000.

[171] B. T. Jap, S. Lai, P. Fisher, and E. Bekiaris, "Using Spectral Analysis to extract frequency components from Electroencephalography: Application for Fatigue Countermeasure in Train Drivers," *Proceedings of the 2nd International Conference on Wireless Broadband and Ultra Wideband Communications*, 2007.

[172] E. N. Bruce, *Biomedical Signal Processing and Signal Modeling*. New York: John Wiley and Sons, 2001.

[173] S. Burke, "Analysis of Variance", RHM Technology Ltd, UK, Statistics and Data Analysis Report, 1998.

[174] E. L. van de Broek, V. Lisy, J. H. Janssen, J. H. D. M. Westerink, M. H. Schut, and K. Tuinenbreijer, "Affective Man-Machine Interface: unveiling Human Emotions through Biosignals," in *Proceedings of the Biostec*, 2010, pp. 21-47.

[175] G. G. Baghdadi and A. M. Nasrabadi, "Comparison of Different EEG Features in Estimation of Hypnosis Susceptibility," *Journal of Computing Biology Medicine*, vol. 42, no 5, pp. 590-597, 2012.

[176] J. D. Bonita, A. M. Albano, L. C. C. Ambolode II, and P. E. Rapp, "Study of Correlation of Multichannel Human EEG," in *Proceedings of the 11th SPVM National Physics Conference and Workshop*, 2009, pp.31-33.

[177] E. C. Ifeachor and B. W. Jervis, *Digital Signal Processing: A Practical Approach.* New Jersey: Prentice Hall, Inc., 2002.

[178] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing: Principles, Algorithms, and Applications*. New Jersey: Prentice Hall, Inc., 1996.

[179] M. Fastenrath, K. J. Friston and S. J. Kiebel, "Dynamical causal modeling for M/EEG: Spatial and temporal symmetry constraints," *Journal of NeuroImage*, vol. 44, pp. 154-163, 2009.

[180] W. Lei, L. Khan and B. Thuraisingham, "An Effective Evidence Theory Based k-Nearest Neighbor (k-NN) Classification," in *Proceedings of the IEEE/WIC/ACM International Conference*, 2008, pp. 797-801.

[181] T. Denoeux, "A k-Nearest Neighbor Classification Rule Based on Dempster-Shafer Theory," *IEEE Transactions on Systems, Man, and Cybernatics*, vol. 25, no. 5, pp. 804-813, 1995.

[182] "Evaluating a classification model". Internet: http://www.compumine.se/web/public/newsletter/20071/precision-recall.html, January, 2012.

[183] "Confusion matrix". Internet: http://www.mathworks.com/matlabcentral /fileexhange//21212-confusion-matrix- matching-matirx/recall.html, January, 2012.

[184] R. Kohavi and Provost, "Special Issue on Applications of Machine Learning and the Knowledge Discovery Process," *Machine Learning*, vol. 30, pp. 271-274, 1998

[185] O. AlZoubi, R. A. Calvo, and R. H. Stevens, "Classification of EEG for Affect Recognition: An Adaptive Approach," in *Proceedings of the International Conference on Artifical Intelligence*, 2009, pp. 52-61.

[186] O. Faust, W. Yu, and N. A. Kadri, "Computer-Based Identification on Normal and Alcoholic EEG Signals using Wavelet Packets and Energy Measures," *Journal of Mechanics in Medicine and Biology*, vol. 13, no 3, pp. 1-17, 2012.

[187] F. C. Galan and C. R. Beal, "EEG Estimates of Engagement and Cognitive Workload Predict Math Problem Solving Outcomes," in *Proceedings of the International Conference on User Modeling, Adaption, and Personalization* 2012, pp. 51-62.

[188] S. R. Yeddula, "Applications of Hilbert Huang Transform (HHT) to the Analysis of EEG and other non-stationary time series," Master Thesis, Lamar University, Texas, United States, 2012.

[189] Z. Ji, T. Sugi, X. Wang and M. Nakamura, "Data Classification for Analysing Characteristics of Electrocardiogram and Electroencephalogram Under Continuous Long-time Mental Calculation and Rest," in *Proceedings of the 2nd Biomedical Engineering and Informatics*, 2009, pp. 1-5.

[190] Bezdek and J. C., *Pattern Recognition with Fuzzy Objective Function Algorithms*. New York: Plenum, 1981.

[191] J. Bezdek†, "Cluster validity with fuzzy sets," *Cybernetics and Systems*, vol. 3, pp. 58-73, 1973.

[192] K. A. Abdul Nazeer, M. P. Sebastian and S. D. Madhu Kumar, "A Heuristic k-Means Algorithm with Better Accuracy and Efficiency for Clustering Health Informatics Data", *Journal of Medical Imaging and Health Informatics*, vol. 1, pp. 66-71, 2011.

[193] G. Baghdadi and A. M. Nasrabadi, "Classifying Hypnotizable Groups Using EEG Weighted Regional Frequency," *Computer Science & Engineering and Electrical Engineering*, vol. 17, no. 1, pp. 71-80, 2010.

[194] L. J. Street, *Introduction to Biomedical Engineering Technology*. Florida: CRC Press, 2008.

[195] C. J. James, O. Gibson, and M. Davies, "On the analysis of single versus multiple channels on electromagnetic brain signals," *Artificial Intelligence in Medicine*, vol. 37, pp. 131, 2006.

[196] M. Mikhail, "Using Minimal Number of Electrodes for Emotion Detection using Noisy EEG data," M. S. Thesis, The American University in Cairo, Cairo, Egypt, 2010.

[197] Z. Krajewski and E. Tkacz, "Power Noise Removal for Monitoring of Biomedical Signals in Daily Life Applications," *Journal of Medical Imaging and Health Informatics*, vol. 2, pp. 44-48, 2012.

[198] H. Schwilden, "Concepts of EEG processing: from power spectrum to bispectrum, fractals, entropies and all that," *Best Practice & Research Clinical Anaesthesiology*, vol. 20, no. 1, pp. 31-48, 2006.

[199] A. Zhang, B. Yang and L. Huang, Feature Extraction of EEG Signals Using Power Spectral Entropy, in *Proceedings of the International Conference on Biomedical Engineering and Informatics*, 2008, pp. 435-438.

[200] R. W. Thatcher, C. J. Biver, and D. M. North, "Z Score EEG Biofeedback: Technical Foundation," NeuroImaging Laboratory, Applied Neuroscience, Inc., Florida, USA, Technical Report, 2007.

[201] G. Kecklund and T. Akerstedt, "Report on Methods and Classification of Stress, Inattention and Emotional States," Information Society Technologies (IST), Sensation, 2004.

[202] C. D. Katsis, N. S. Katersidis and D. I. Fotiadis, "An Integrated System based on Physiological signals for the assessment of affective states in patients with anxiety disorders", *Journal of Biomedical Signal Processing and Control*, pp.1-8, 2010.

[203] E. Jovanov, A. O. Lords, D. Raskovic, P. Cox, R. Adhami and F.Andrasik, "Stress Monitoring Using a Distributed Wireless Intelligent Sensor System", *Engineering in Medicine and Biology Magazine, IEEE*, vol. 22, no. 3, pp. 49-55, 2003.

.

APPENDICES

APPENDIX A

Human Subject Demographic Data

FORM A

Human Subject Demographic Data
Health Declaration and Medical Examination Form

Biomedical Research And Development Laboratory For Human Potential
Faculty Of Electrical Engineering UiTM Shah Alam 40450 Tel: 55436002

Subject Reference Number: (for research use only) 2 **Date:** 24/3/09

A	Demographic Data	☑ UiTM Student ☐ UiTM Staff ☐ Others	
1	Name	ZODIE MOHAMED HANAFIAH	
2	IC Number (new/old)	830910-14-6029	
3	Date of birth	10 September 1983	
4	Profession/Occupation	STUDENT	
5	Age	25	
6	Race	☑ Malay	☐ Non Malay
7	Religion	☑ Muslim	☐ Non Muslim
8	Gender	☑ Male	☐ Female
9	Marital Status	☑ Single	☐ Married
10	Telephone Number (H/p, office, home)	016-2003269	
11	Mailing Address	—	
12	Email Address	sb8zee_10@yahoo.com	
13	Name, relationship and contact number of next of kin (in case of emergency)		
B	Health Declaration		
1	Pregnancy Status	☐ Pregnant ☑ not pregnant	
2	Are you taking any form of medication/drugs?	☑ No ☐ Yes ________ (please specify)	
3	Present Health condition (In general)	☑ Healthy ☐ Unhealthy ________ (please specify)	
4	Past Health condition (In general)	☑ Healthy ☐ Unhealthy ________ (please specify)	

5	Have You ever suffered any of the following conditions	Yes	No
i	Psychiatric Illness (Sakit Jiwa)		✓
ii	Epilepsy (Sawan)		✓
iii	Migraine (Migraine)		✓
iv	Hysteria (Histeria)		✓
v	Allergic Rhinitis (Resdung)		✓
vi	Asthma (Lelah)		✓
vii	Tuberculosis (Batuk Kering)		✓
viii	Hypertension (Darah Tinggi)		✓
ix	Diabetes Mellitus (Kencing Manis)		✓
x	Heart Diseases (Penyakit Jantung)		✓
xi	Thyroid (Tiroid)		✓
xii	Kidney Diseases (Penyakit Buah Pinggang)		✓
xiii	Gastric (Gastrik)		✓
xiv	HIV/AIDS		✓
xv	Cancer		✓
xvi	Venereal Diseases (Penyakit Kelamin)		✓
xvii	Leukaemia		✓
xviii	Hepatitis		✓
xix	Other illnesses (Penyakit Penyakit Lain)		✓
xx	Surgical (Pembedahan)		✓
xxi	Allergic (Alahan)		✓
xxii	Disability (Kecacatan)		✓

C	Medical Examination	Carried out by a qualified medical officer
1	Weight (kilogram)	
2	Height (meter)	
3	Blood Pressure	
4	Pulse Rate	
5	Body Mass Index	
6	Are there abnormalities of the following system? If yes, please describe.	

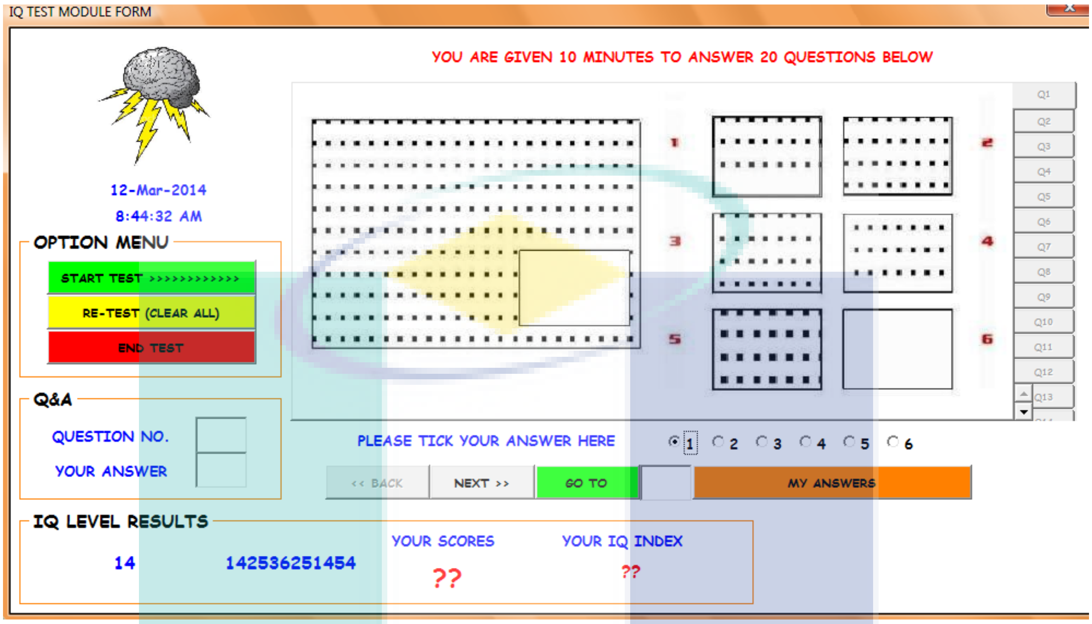
	Systems	Normal	Abnormal	Comment
i	Skin	✓		
ii	Head	✓		
iii	Eyes	✓		
iv	Ears	✓		
v	Nose	✓		
vi	Mouth	✓		
vii	Neck	✓		
viii	Chest	✓		
ix	Breasts	✓		
x	Cardiovascular	✓		
xi	Abdomen	✓		
xii	Neurology	✓		

7	Results of the medical examination	☐ **Medically Fit** ☐ **Limited Capability** ☐ **Unfit**
8	(medical officer's signature) Date:	Name/Post/Qualification/Stamp
9	(subject's signature) Date: 24/3/09	Declaration: I hereby, voluntarily agree to participate in this research and the information given above is correct up to this date.
10	(investigator's signature) Date: 24/3/09	NORIZAM BIN SULAIMAN LECTURER FACULTY OF ELECTRICAL & ELECTRONICS ENGINEERING UNIVERSITI MALAYSIA PAHANG 26600 PEKAN PAHANG DARUL MAKMUR Name/Post/Qualification/Stamp

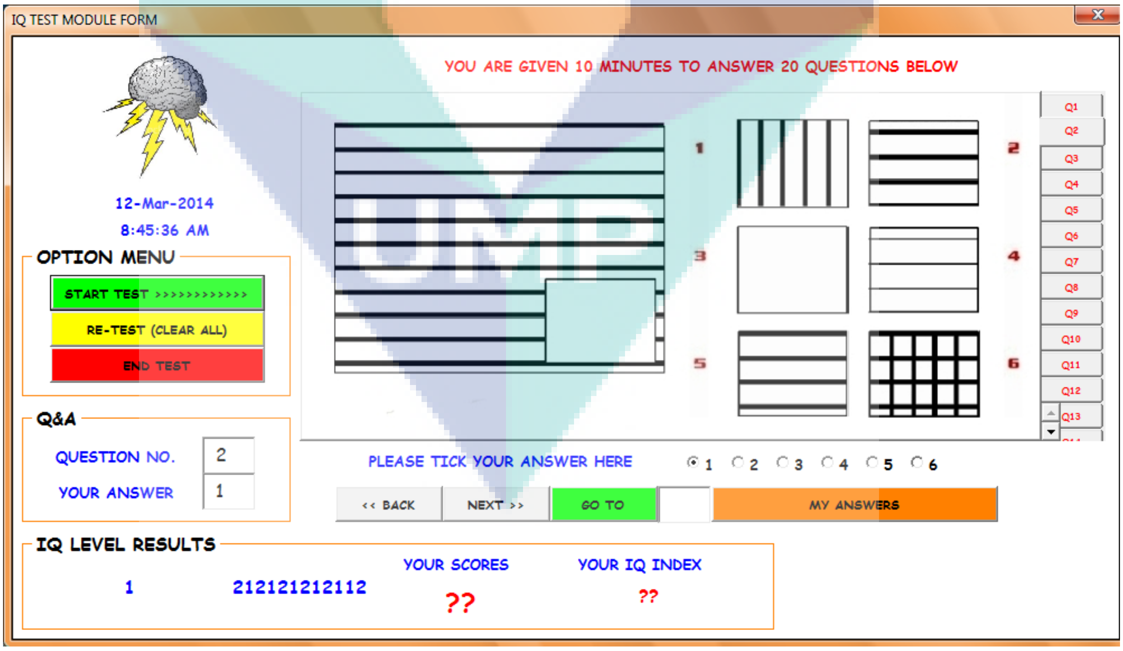
APPENDIX B

Example of IQ Test Questionnaires – Non-Verbal Assessment

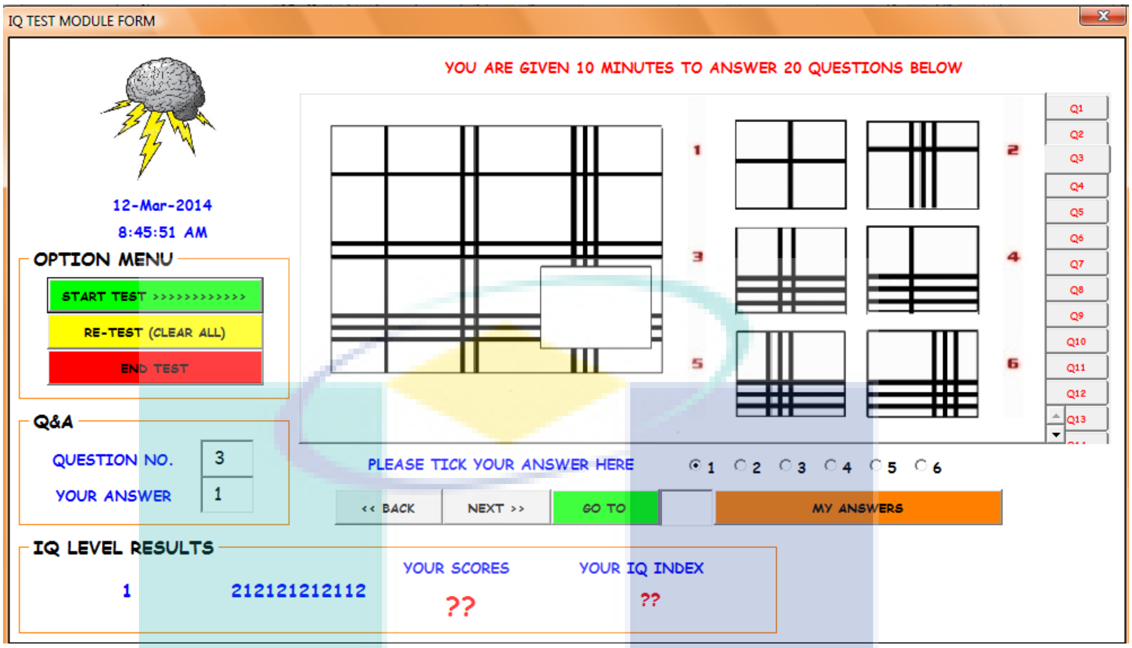
Question 1:



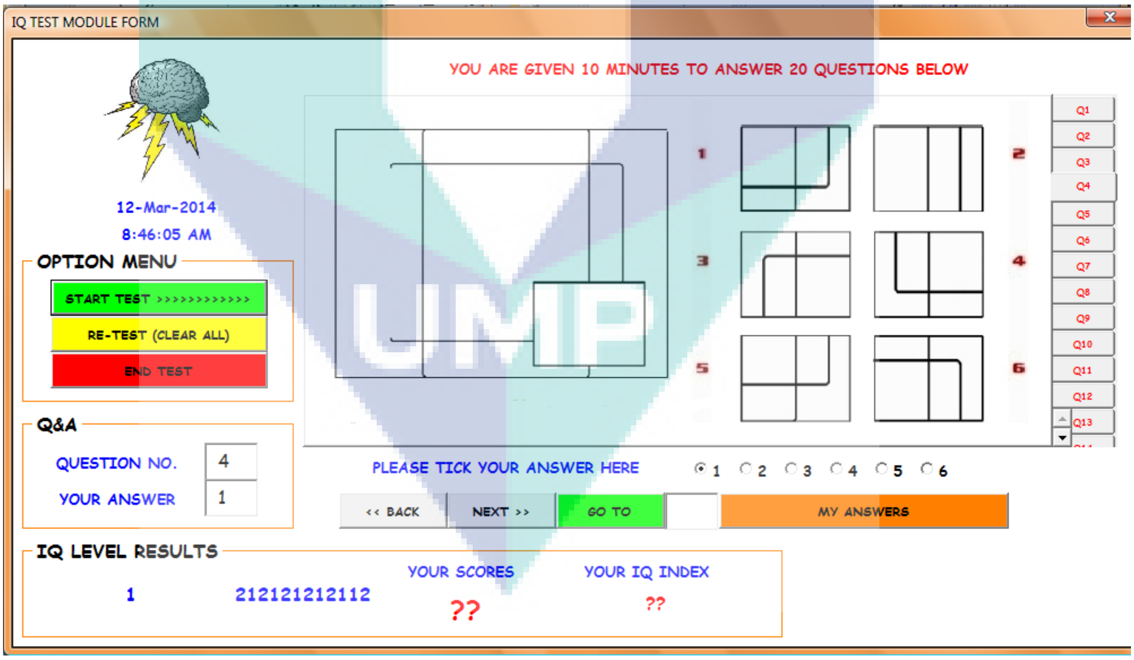
Question 2:



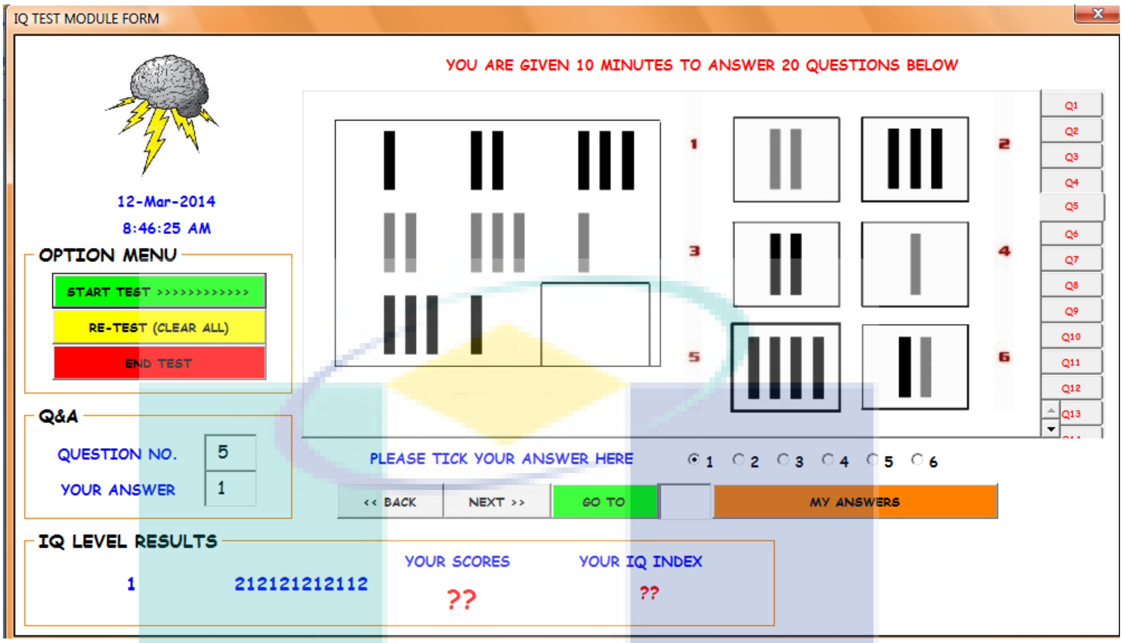
Question 3:



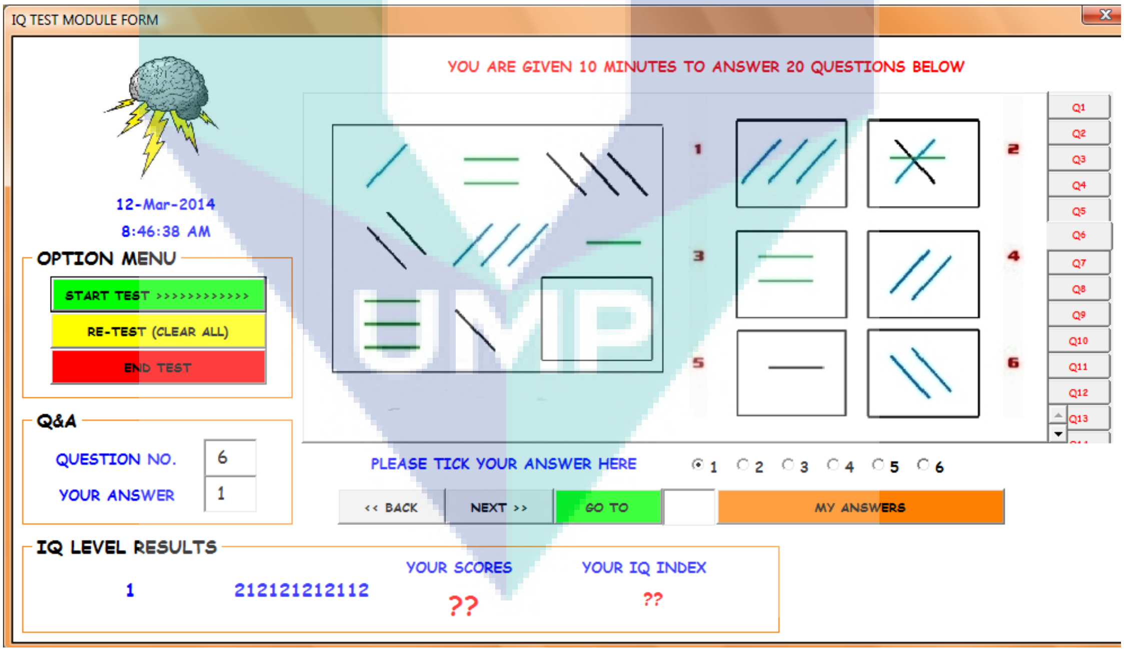
Question 4:



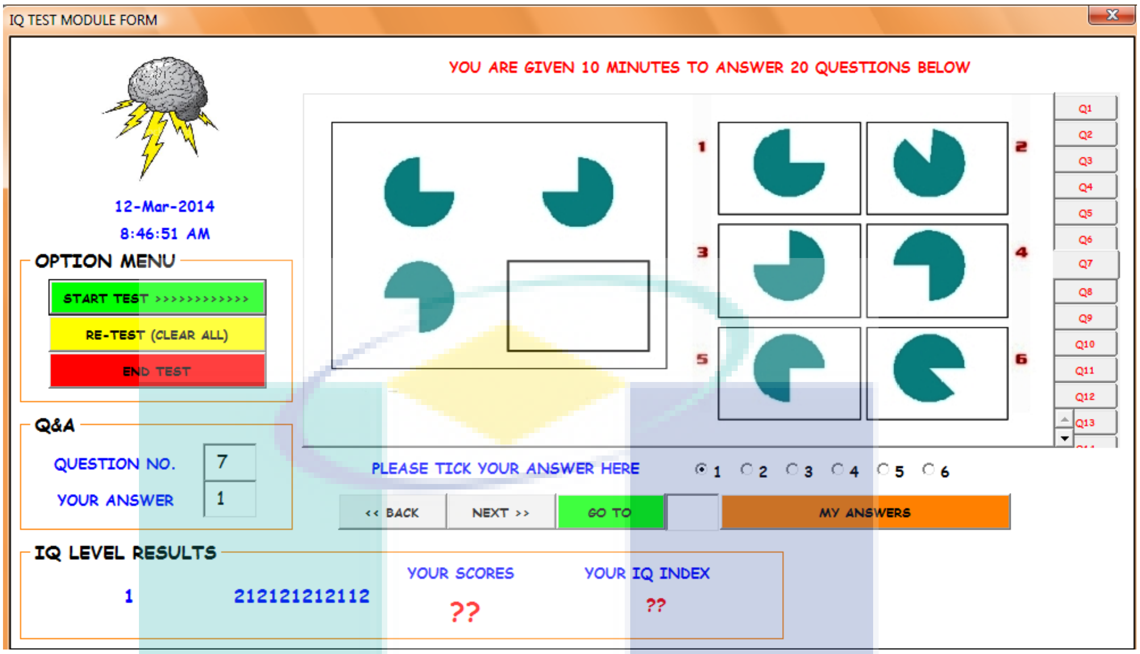
Question 5:



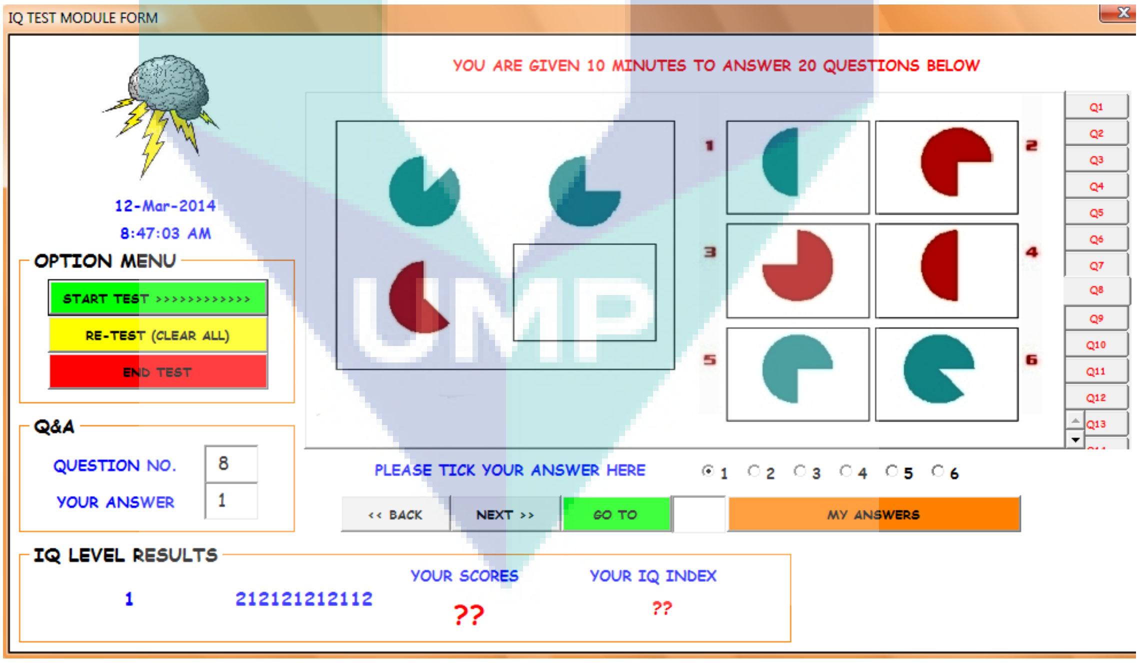
Question 6:



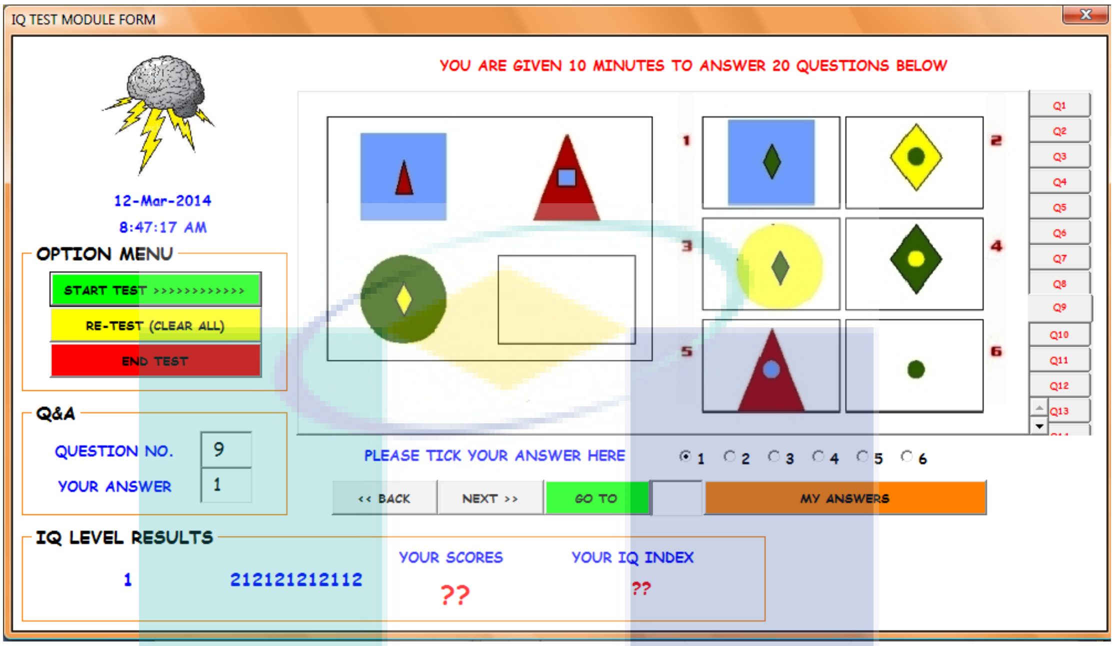
Question 7:



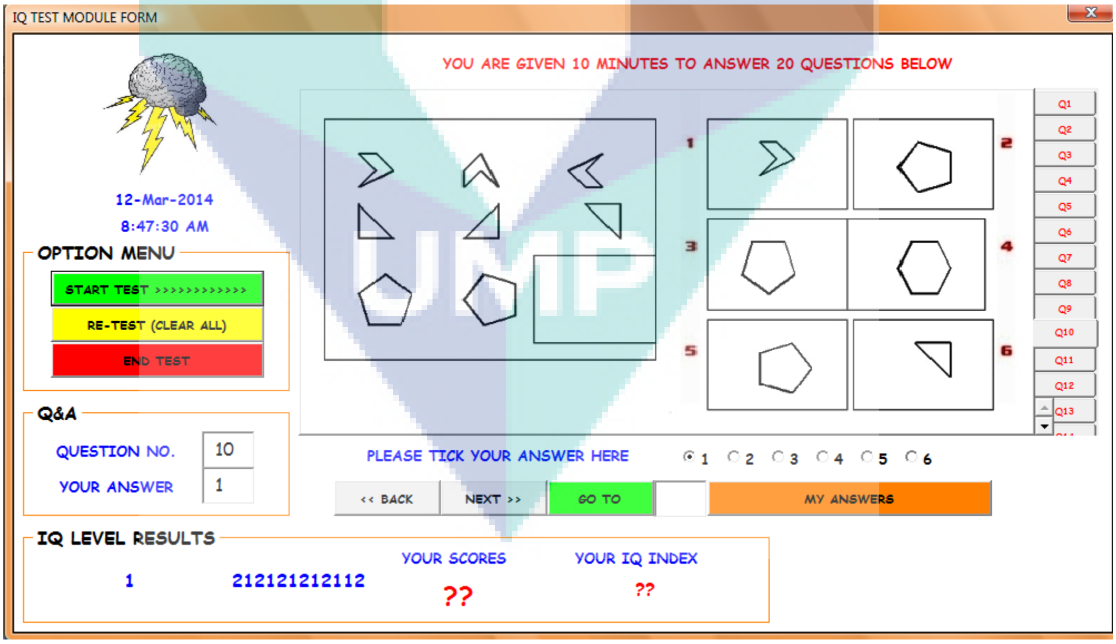
Question 8:



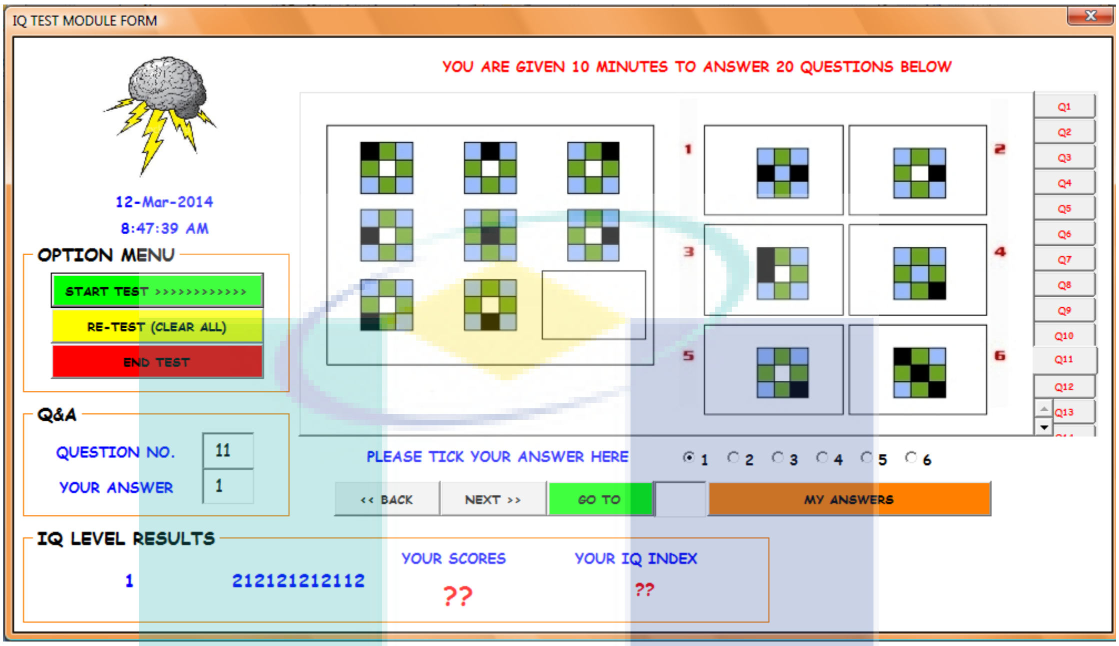
Question 9:



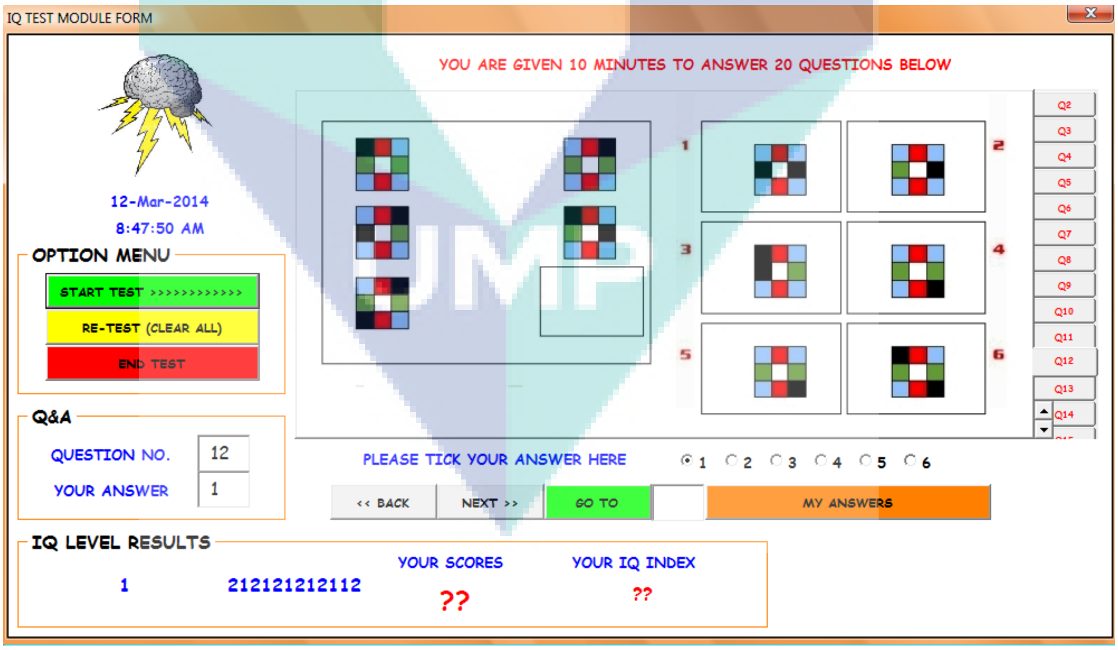
Question 10:



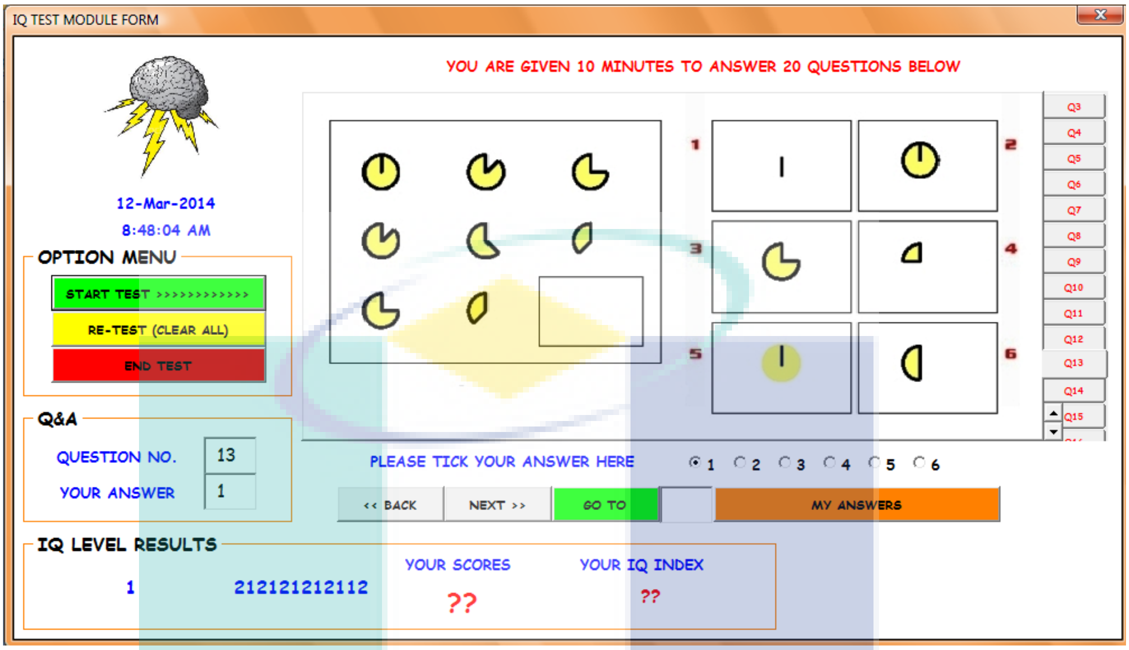
Question 11:



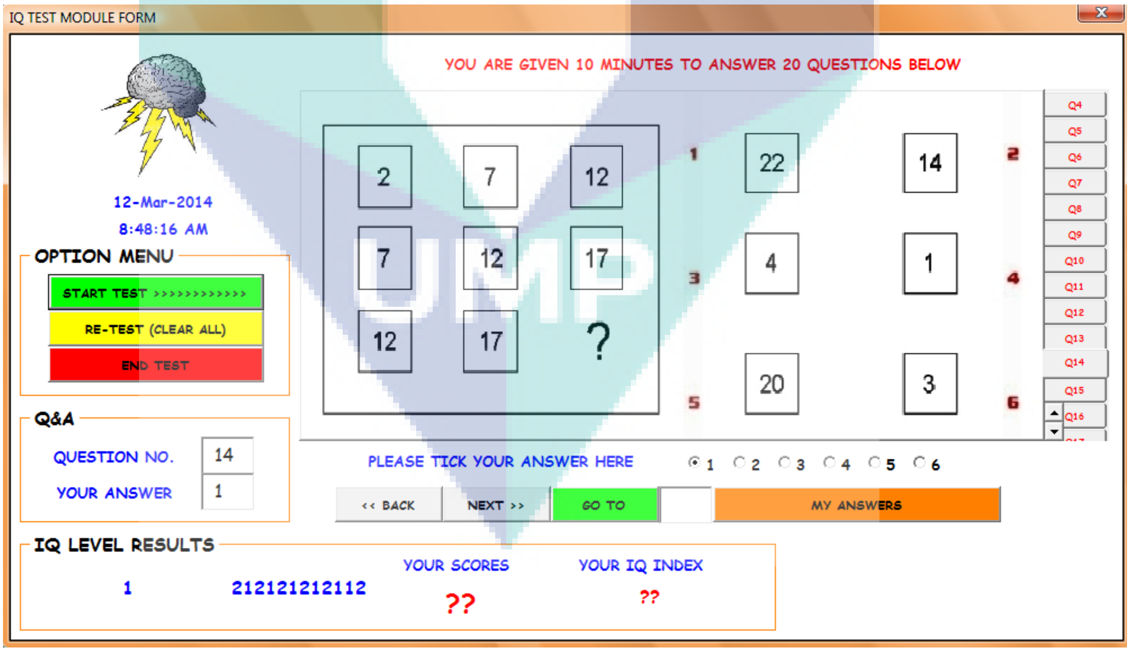
Question 12:



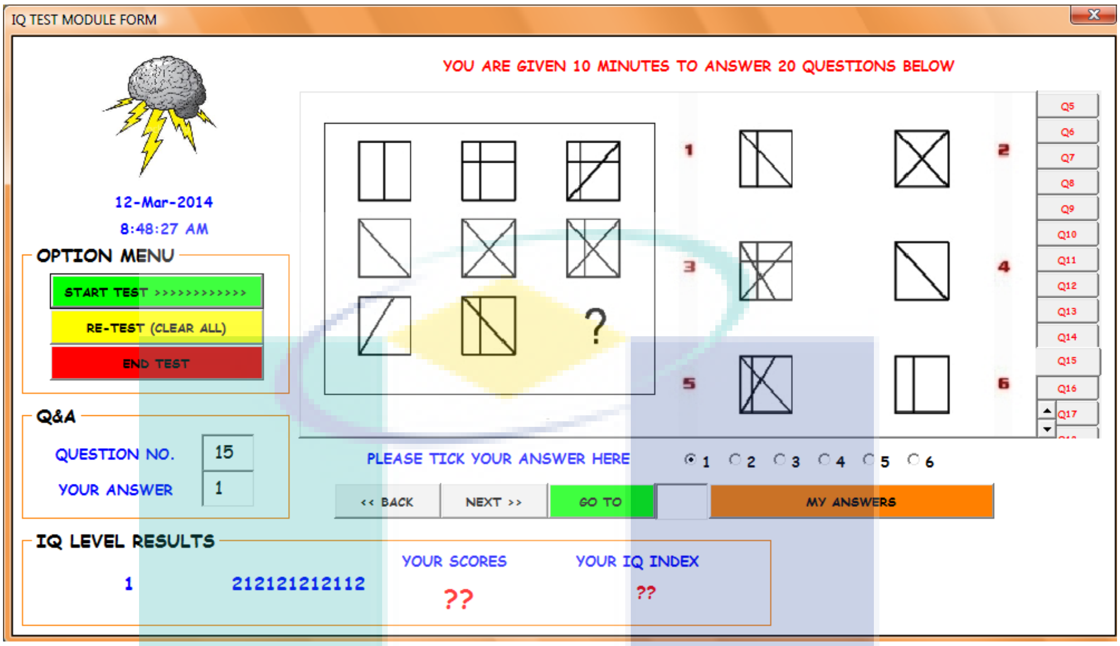
Question 13:



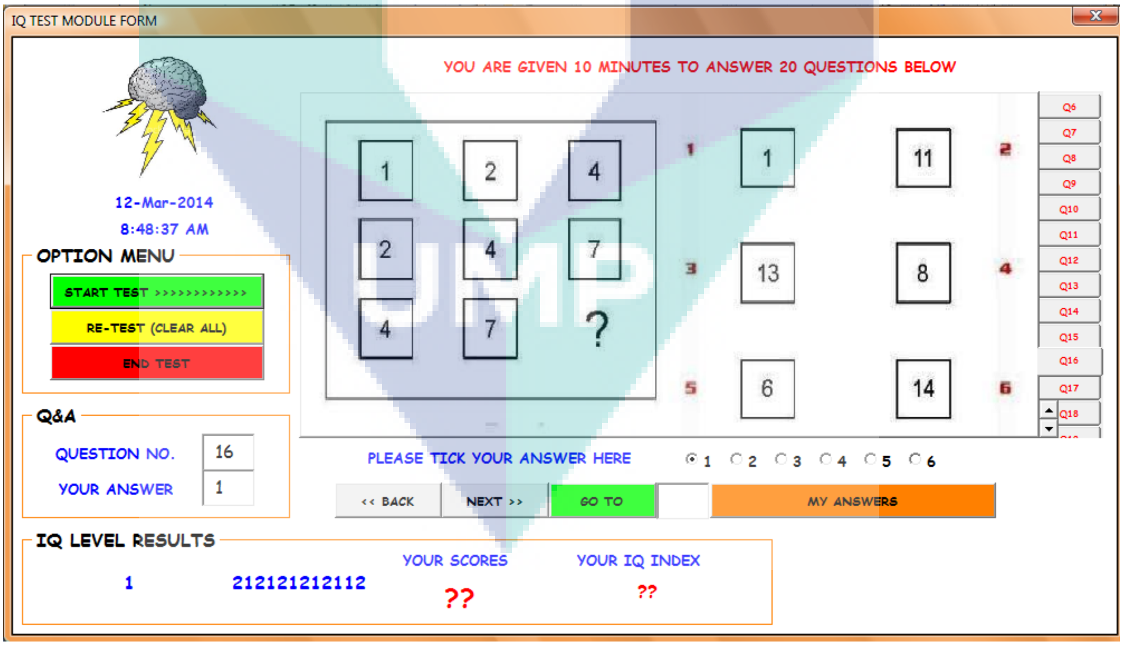
Question 14:



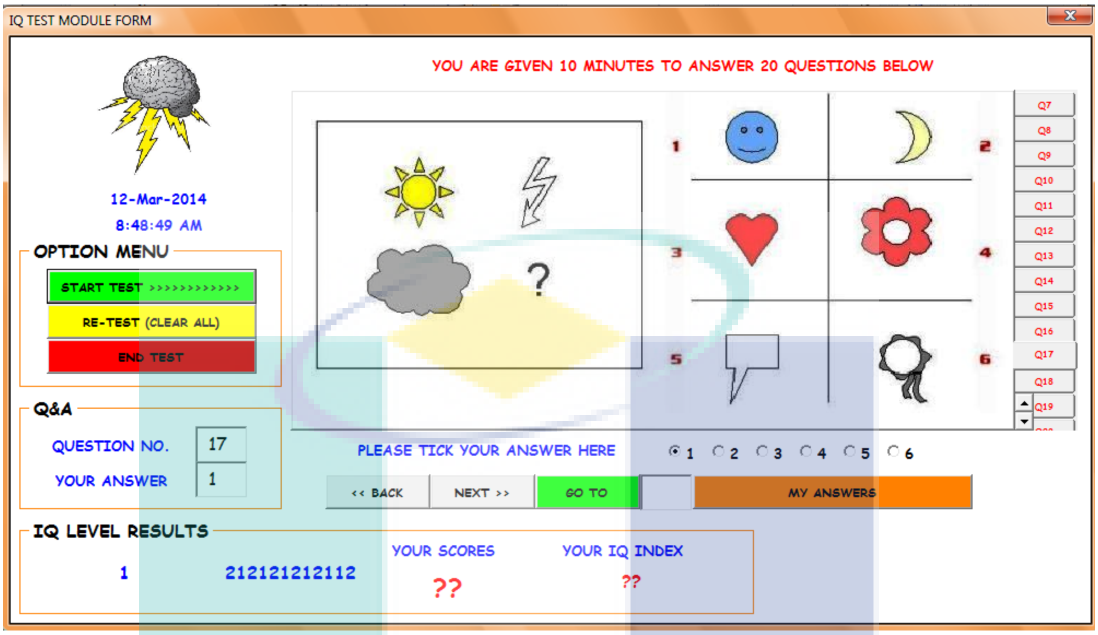
Question 15:



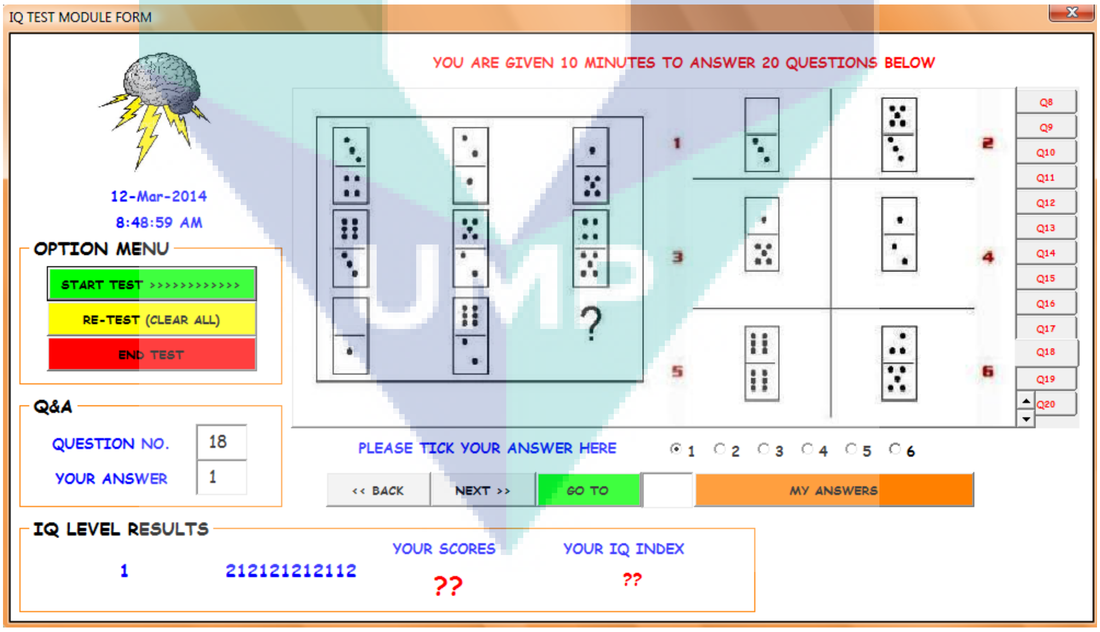
Question 16:



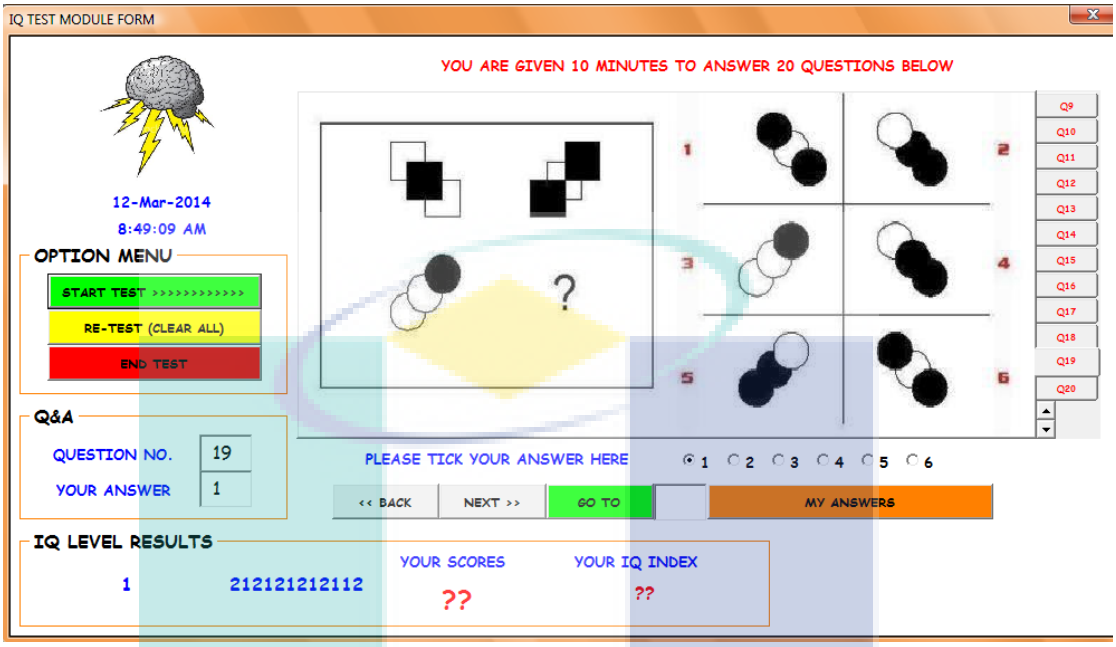
Question 17:



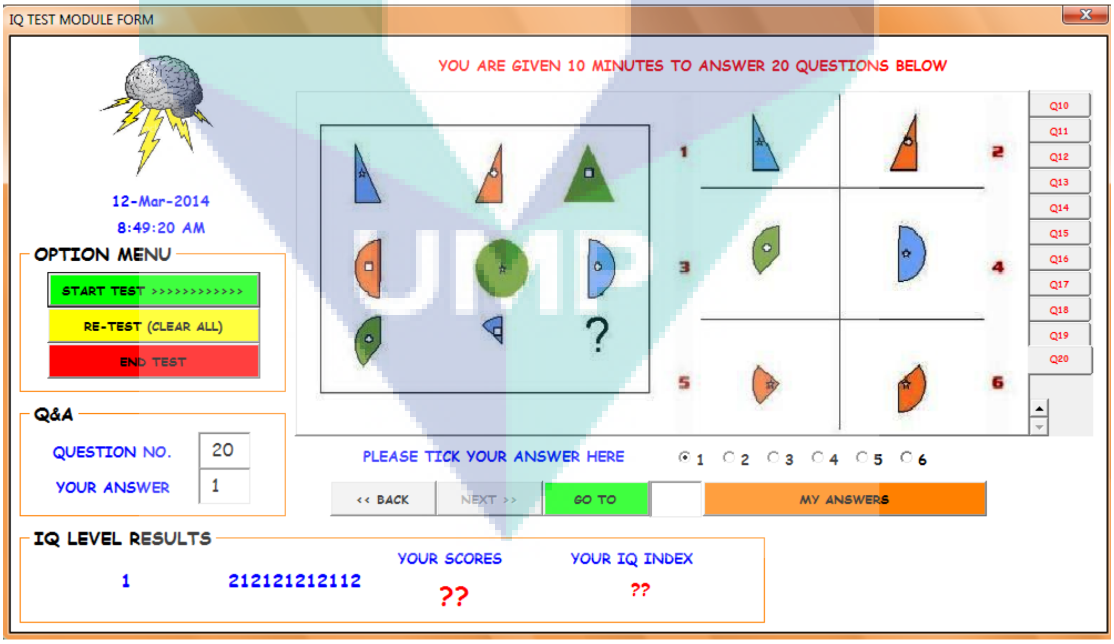
Question 18:



Question 19:



Question 20:



APPENDIX C

1. EEG Energy Spectral Density Across Frequency Bands for Group 1

Subject	Delta_R	Delta_L	Theta_R	Theta_L	Alpha_R	Alpha_L	Beta_R	Beta_L
1	1618.47	1882.19	491.75	556.20	287.86	321.97	48.02	61.78
2	2231.18	3136.05	474.52	493.69	105.86	89.38	46.06	34.70
3	2773.13	2490.88	486.19	435.95	220.89	202.13	100.59	94.47
4	15824.54	10467.58	984.34	819.27	242.02	221.05	66.20	66.68
5	2831.32	5749.47	903.03	1416.69	775.43	940.27	68.00	100.59
6	19718.88	18917.70	2272.11	2158.44	616.71	637.93	71.23	88.38
7	8007.66	7858.84	1562.72	1545.04	609.16	581.96	86.11	72.73
8	10148.42	5174.07	1230.09	974.32	1083.59	913.24	89.66	91.69
9	6826.18	4377.86	759.30	799.18	156.75	151.08	41.37	42.75
10	1141.56	1496.53	715.73	678.20	1208.01	846.99	147.89	131.23
11	19920.63	11745.43	1289.05	1139.61	443.26	416.23	68.74	62.89
12	2457.13	5136.64	775.83	1032.80	472.76	586.19	80.08	85.59
13	9670.18	9285.50	1443.93	1441.63	422.94	462.42	114.77	115.32
14	9753.67	8583.19	1850.53	1683.97	639.77	604.75	63.50	63.53
15	3579.32	2745.33	906.11	693.13	569.16	434.17	81.89	93.39
16	2720.52	1747.57	644.74	473.02	334.59	330.58	45.52	53.24
17	3156.58	2589.01	536.01	710.24	247.36	328.40	32.68	34.93
18	1712.52	2172.49	692.09	791.15	1373.65	1420.97	58.24	63.11
19	7343.46	9387.98	1550.52	1903.78	325.09	405.39	36.96	48.37
20	4319.01	5253.46	720.62	937.67	502.89	530.87	37.43	37.01
21	5629.64	8256.46	1468.24	1617.84	1051.36	902.88	73.80	61.70
22	4634.11	3570.46	437.61	621.87	80.45	136.68	36.83	46.20
23	4825.55	2978.56	612.00	462.99	396.94	389.47	46.44	55.70
24	20367.76	19237.78	4614.00	3693.95	937.35	626.45	104.32	71.66
25	4088.28	4205.23	883.15	889.50	785.14	713.76	99.15	88.88
26	3059.17	2314.77	675.24	594.47	408.57	360.84	42.46	50.91
27	3899.05	3177.60	738.36	842.99	502.06	619.80	109.67	125.53
28	2874.56	2510.90	1051.76	978.51	2175.57	2105.68	105.96	104.03
29	13883.16	12864.18	1237.47	1123.10	400.12	360.35	76.22	78.67
30	1783.72	2486.09	944.77	767.77	924.37	706.64	61.16	61.51
31	8271.91	7304.96	1347.41	1357.08	1130.80	1322.13	70.80	76.24
32	3965.14	5348.45	1062.69	1043.51	885.71	725.42	73.07	74.56
33	3919.85	4785.80	512.11	556.42	181.57	165.04	49.92	52.84
34	4285.02	4028.12	648.20	623.56	303.64	318.68	58.36	49.73
35	5238.72	3040.23	693.11	614.36	373.54	425.55	42.55	59.64
36	5551.34	9972.65	817.90	1179.53	699.56	721.35	226.76	164.21
37	7376.82	4337.33	1499.60	1119.94	599.90	511.12	83.50	63.52
38	6632.90	6856.55	1110.19	1286.89	708.31	703.65	39.17	43.33
39	13219.73	16136.94	1000.32	2530.05	207.95	539.09	68.91	142.71

40	2537.49	2819.50	474.10	593.61	301.38	332.42	127.19	125.17
41	3179.76	5199.47	539.94	726.07	482.99	557.10	47.07	43.37
42	4045.63	10809.52	1440.35	2192.67	1353.08	1371.77	51.55	54.72
43	2806.92	5752.25	899.88	1423.35	782.73	945.46	69.50	102.45
44	7334.46	9402.30	1550.69	1908.49	325.40	403.03	38.93	50.45
45	4646.60	3616.35	440.37	625.36	82.56	138.51	38.58	48.46
46	20423.85	19304.02	4627.89	3706.87	941.83	629.94	106.14	74.33
47	5242.16	3043.33	694.76	614.23	376.42	427.75	44.44	61.95
48	5555.06	9953.09	821.67	1185.21	702.78	722.70	228.68	164.88
49	7383.20	4335.49	1502.05	1122.67	602.64	509.91	85.29	65.05
50	13219.89	16103.30	1001.33	2541.57	208.76	544.68	70.50	145.34

2. EEG Energy Spectral Density Across Frequency Bands for Group 2

Subject	Delta_R	Delta_L	Theta_R	Theta_L	Alpha_R	Alpha_L	Beta_R	Beta_L
1	10206.90	7942.12	1890.31	1617.02	564.67	496.70	384.16	303.82
2	25824.87	23602.12	5976.11	5420.75	1528.56	1428.45	349.62	297.47
3	7957.24	6471.60	1580.71	1346.68	350.20	416.70	109.19	176.20
4	11715.45	7847.78	2538.35	1850.65	550.04	383.18	575.74	183.97
5	13725.09	17980.57	3383.51	4143.68	747.99	822.15	479.08	422.08
6	8010.86	7177.60	1880.11	1684.32	341.59	326.38	84.67	90.53
7	9564.84	8401.29	2039.69	1968.67	398.19	421.18	97.03	125.45
8	6073.68	6442.81	1149.19	1319.47	267.05	310.03	92.60	108.68
9	5360.78	4842.79	1318.82	1222.29	387.47	433.35	178.09	252.99
10	21996.95	16554.16	4875.73	4219.34	1177.16	1075.42	802.00	793.60
11	13930.89	12808.59	2280.16	2263.39	554.72	609.73	270.25	449.87
12	14058.72	12611.46	3197.90	2962.65	1099.55	1214.72	234.13	173.18
13	5491.67	6444.69	1224.68	1139.73	285.81	263.37	66.69	65.63
14	6577.91	7039.22	1766.76	2046.02	364.21	401.14	97.77	85.76
15	26486.44	28533.36	6477.70	6772.10	1369.96	1505.31	189.59	224.43
16	11435.20	12834.18	3149.44	3950.40	695.40	795.11	107.31	103.62
17	12981.80	12144.45	2801.76	2597.48	639.22	624.86	159.93	189.41
18	8374.07	9830.59	1952.74	2443.03	364.64	437.77	78.23	73.33
19	12262.37	13094.96	2045.77	2261.71	423.15	435.81	111.48	100.93
20	9536.36	7023.68	1943.27	1663.28	401.93	390.28	88.95	101.65
21	16207.15	21560.77	3231.42	4351.95	662.72	868.47	171.89	175.83
22	12399.96	11909.80	1586.94	1811.09	370.77	405.39	190.79	108.60
23	5363.04	4943.64	1146.52	1013.25	264.20	225.53	90.64	124.11
24	10023.99	7400.09	1921.46	1550.69	394.89	323.44	120.00	119.98
25	7246.18	5942.19	1481.10	1182.54	274.85	233.94	68.50	64.09
26	5572.87	6952.59	1047.89	984.01	163.10	139.47	101.95	62.69
27	13092.22	12513.42	2272.24	2213.73	466.72	460.56	107.70	117.02

28	9228.59	11488.58	2045.21	2342.59	465.77	496.41	136.93	219.27
29	13181.27	14152.85	2641.65	2851.57	471.93	493.03	121.55	106.19
30	10389.74	17031.07	2089.40	2841.39	431.83	534.32	130.24	124.60
31	10212.65	12598.95	2353.92	4351.17	604.51	1314.25	411.53	573.90
32	16075.55	13073.72	4469.26	3630.80	924.62	750.41	293.80	185.55
33	17643.95	15627.80	3339.50	2953.03	732.36	657.70	175.83	159.52
34	27873.14	31035.08	5403.91	5970.54	888.87	934.50	216.43	227.00
35	22944.28	13382.91	4818.05	2788.17	987.54	498.69	270.29	143.34
36	14208.82	21319.63	2858.68	4578.92	674.43	1100.48	104.65	212.93
37	20801.06	12250.92	4679.41	2412.94	1045.20	517.75	367.86	161.21
38	20567.00	29234.34	5927.21	9031.21	1583.40	2208.87	231.94	305.31
39	10020.11	6969.38	2316.58	1570.54	497.84	369.81	118.27	277.98
40	21971.92	19780.97	4441.23	3855.73	857.80	725.61	172.52	146.63
41	19299.50	18796.48	5247.92	3968.97	1302.53	729.76	209.63	113.08
42	17828.93	14932.07	3202.41	2814.04	614.45	512.89	110.83	107.17
43	7410.39	7663.51	1802.64	1861.65	398.20	394.77	140.70	127.25
44	9973.64	11664.96	2195.34	2397.61	479.78	529.32	136.26	341.35
45	9608.48	10054.22	2130.51	2247.21	738.40	902.21	592.81	1182.41
46	6367.99	4250.47	1178.00	901.06	268.29	199.24	113.62	50.51
47	32555.94	15049.54	6393.62	3621.10	1119.65	671.80	172.27	134.93
48	11829.16	7813.02	2541.33	1850.05	541.20	376.05	561.44	191.10
49	9370.02	11374.06	2067.09	2337.87	475.54	504.59	140.55	220.40
50	6814.23	4879.71	1284.49	1097.39	293.21	254.53	118.60	65.23

3. EEG Energy Spectral Density Across Frequency Bands for Group 3

Subject	Delta_R	Delta_L	Theta_R	Theta_L	Alpha_R	Alpha_L	Beta_R	Beta_L
1	6833.62	2600.97	511.56	458.79	273.43	300.54	63.60	70.95
2	5933.96	10810.80	867.00	1496.21	225.85	314.93	82.28	90.33
3	6317.49	5952.59	928.83	952.65	310.52	257.19	72.46	63.87
4	5143.10	3386.42	866.62	658.50	252.83	290.27	65.39	60.17
5	4099.96	2418.17	1140.67	1275.88	1208.42	1808.85	225.03	453.16
6	2493.82	2893.84	603.62	602.35	325.76	333.43	73.27	80.68
7	1480.99	2443.67	916.23	1052.42	764.84	896.74	60.71	115.43
8	3389.62	3598.15	517.29	633.46	214.93	229.22	64.54	60.39
9	8461.04	6822.72	1181.80	1020.44	268.19	260.72	43.92	50.89
10	21955.19	14925.21	4241.45	2457.01	1007.08	672.33	140.51	90.27
11	6943.64	15058.75	1067.66	1542.51	191.86	205.46	39.36	37.98
12	8910.71	10413.40	1361.68	1550.39	441.52	446.76	51.61	46.02
13	6643.71	5810.36	807.48	678.57	599.89	576.85	70.31	45.27
14	17612.69	18836.47	3954.38	4744.28	1215.77	1280.81	131.94	139.99
15	2344.59	3142.66	1150.41	1236.00	960.32	860.30	69.31	75.87
16	5798.05	6765.13	1162.41	1123.23	401.03	355.80	105.47	105.34
17	1977.39	1976.37	606.00	525.67	657.27	495.09	123.47	99.00
18	1234.44	1714.47	304.27	322.66	209.30	209.93	23.77	31.33
19	2997.69	4329.07	795.52	942.31	523.60	564.02	62.38	66.22
20	13113.07	12467.68	2816.10	2697.51	2444.91	2472.07	151.42	81.70
21	14489.86	17348.12	1470.19	1643.48	212.87	240.53	48.81	62.85
22	2860.04	2739.90	470.73	420.10	104.21	100.92	30.41	31.78
23	7707.98	9762.32	1649.79	1580.60	330.19	628.86	89.52	95.21
24	11550.11	9651.30	1853.93	1653.71	566.13	455.72	92.80	61.08
25	3073.15	3578.25	559.83	688.53	272.40	331.31	40.96	58.59
26	10814.19	10248.18	1711.49	1799.00	188.88	220.72	54.72	48.91
27	2909.07	4037.85	558.54	611.47	188.55	177.39	19.70	20.91
28	2397.83	3649.20	490.02	868.57	638.77	848.93	57.07	124.17
29	1234.44	1714.47	304.27	322.66	209.30	209.93	23.77	31.33
30	6828.94	2601.53	513.06	461.21	274.88	302.84	65.01	72.74
31	5913.14	10786.76	869.29	1497.53	227.41	316.49	83.56	92.87
32	5150.35	3386.97	869.75	660.75	254.77	292.58	67.77	61.85
33	9428.42	8567.82	1430.76	1691.55	691.82	822.69	80.50	94.40
34	8470.56	6808.55	1183.87	1025.82	270.63	263.02	45.90	52.66
35	21978.84	14800.64	4220.47	2460.84	1001.51	679.02	143.47	92.96
36	17777.25	18931.44	3968.65	4766.11	1220.41	1289.90	133.61	141.16
37	1976.79	1979.59	609.73	531.67	660.02	499.43	125.05	100.99
38	2995.56	4336.21	799.27	945.92	525.88	566.19	64.65	68.48
39	2856.82	2735.43	474.12	422.24	106.98	103.27	32.12	33.57
40	2910.27	4034.26	563.62	614.02	192.05	180.76	21.50	22.57

4. EEG Energy Spectral Density Across Frequency Bands for Group 4

Subject	Delta_R	Delta_L	Theta_R	Theta_L	Alpha_R	Alpha_L	Beta_R	Beta_L
1	2536.47	2373.52	480.71	732.20	453.99	527.28	87.09	119.73
2	8495.61	5617.62	1131.90	1002.20	254.08	298.89	76.38	77.64
3	9238.77	17031.21	1700.18	3089.75	386.52	706.12	124.87	187.06
4	1631.86	1925.13	617.25	702.20	191.93	219.82	38.39	39.57
5	19566.36	17304.76	4211.93	3699.38	955.03	808.37	105.80	110.11
6	2817.33	2575.47	520.91	484.83	289.36	284.86	71.76	70.69
7	24286.47	29577.74	3325.16	4838.71	623.70	821.11	82.32	104.53
8	4579.85	3645.63	1635.91	1476.28	962.43	989.66	56.12	53.03
9	2994.12	3850.07	469.41	569.82	318.12	346.08	52.86	82.53
10	4569.33	3330.57	737.32	644.97	206.34	193.98	49.01	47.25
11	11758.42	10639.25	2292.67	2072.77	667.17	604.46	79.63	99.76
12	36116.25	29284.58	4691.99	4154.16	624.13	560.36	111.60	120.98
13	3791.65	2995.82	595.11	662.10	365.40	368.04	56.76	47.05
14	2144.19	1860.72	434.61	355.85	151.51	133.49	35.75	31.10
15	10961.58	8202.66	2837.98	2122.56	2155.14	1444.05	166.29	142.02
16	2582.38	3597.38	913.92	1033.50	1047.99	940.22	60.47	54.82
17	2567.53	2524.93	625.72	646.84	323.50	337.73	66.87	78.58
18	1560.57	1313.44	503.25	423.56	541.55	401.40	113.41	85.64
19	901.08	850.75	397.79	392.42	451.38	433.61	46.36	53.16
20	3408.33	4806.03	1598.30	1666.17	1042.39	1014.80	70.39	68.94
21	11058.17	13866.99	2844.98	2937.99	2395.33	2498.87	125.70	84.79
22	2483.16	1695.53	452.03	375.94	113.85	110.29	35.24	37.42
23	2003.99	2116.17	315.51	355.94	83.56	103.62	34.58	39.00
24	7393.93	5077.15	984.75	957.74	276.95	285.37	83.98	76.40
25	19149.30	16138.09	4097.87	3561.21	972.76	743.63	109.51	78.37
26	1496.19	1824.47	380.97	432.95	161.21	188.98	34.01	48.04
27	10572.83	9075.25	2418.46	2043.82	244.00	212.50	61.06	38.25
28	2022.72	2623.90	636.65	526.93	192.64	163.03	40.14	35.51
29	1988.26	1358.41	616.84	540.71	1092.51	1230.24	75.69	117.42
30	2542.35	2374.71	483.66	733.34	456.45	529.58	89.06	121.72
31	8479.25	5613.80	1132.80	1004.70	256.66	300.67	78.40	79.82
32	19597.17	17311.15	4232.42	3708.62	959.48	815.75	108.29	112.33
33	23918.38	29330.58	3285.39	4779.63	627.26	815.66	84.29	105.77
34	4578.21	3326.10	740.19	643.79	209.64	195.83	51.62	48.87
35	11772.51	10665.71	2315.47	2088.78	679.13	603.25	80.64	102.79
36	10924.93	8226.00	2839.70	2135.64	2160.68	1447.80	167.60	26.64
37	1562.99	1318.70	504.19	426.71	544.05	402.75	115.19	87.44
38	3401.49	4788.99	1604.28	1667.08	1047.87	1016.97	72.03	70.73
39	1997.46	2114.61	319.30	357.61	86.76	104.63	36.90	40.40
40	10571.38	9063.61	2415.67	2042.77	245.42	214.14	62.46	39.84

5. EEG Features – Normalization of EEG Asymmetry Ratio for Group 1

Subject	Asym_Delta	Asym_Theta	Asym_Alpha	Asym_Beta
1	-0.08	-0.06	-0.06	-0.13
2	-0.17	-0.02	0.08	0.14
3	0.05	0.05	0.04	0.03
4	0.20	0.09	0.05	0.00
5	-0.34	-0.22	-0.10	-0.19
6	0.02	0.03	-0.02	-0.11
7	0.01	0.01	0.02	0.08
8	0.32	0.12	0.09	-0.01
9	0.22	-0.03	0.02	-0.02
10	-0.13	0.03	0.18	0.06
11	0.26	0.06	0.03	0.04
12	-0.35	-0.14	-0.11	-0.03
13	0.02	0.00	-0.04	0.00
14	0.06	0.05	0.03	0.00
15	0.13	0.13	0.13	-0.07
16	0.22	0.15	0.01	-0.08
17	0.10	-0.14	-0.14	-0.03
18	-0.12	-0.07	-0.02	-0.04
19	-0.12	-0.10	-0.11	-0.13
20	-0.10	-0.13	-0.03	0.01
21	-0.19	-0.05	0.08	0.09
22	0.13	-0.17	-0.26	-0.11
23	0.24	0.14	0.01	-0.09
24	0.03	0.11	0.20	0.19
25	-0.01	0.00	0.05	0.05
26	0.14	0.06	0.06	-0.09
27	0.10	-0.07	-0.10	-0.07
28	0.07	0.04	0.02	0.01
29	0.04	0.05	0.05	-0.02
30	-0.16	0.10	0.13	0.00
31	0.06	0.00	-0.08	-0.04
32	-0.15	0.01	0.10	-0.01
33	-0.10	-0.04	0.05	-0.03
34	0.03	0.02	-0.02	0.08
35	0.27	0.06	-0.07	-0.17
36	-0.28	-0.18	-0.02	0.16
37	0.26	0.14	0.08	0.14
38	-0.02	-0.07	0.00	-0.05
39	-0.10	-0.43	-0.44	-0.35
40	-0.05	-0.11	-0.05	0.01
41	-0.24	-0.15	-0.07	0.04

42	-0.46	-0.21	-0.01	-0.03
43	-0.34	-0.23	-0.09	-0.19
44	-0.12	-0.10	-0.11	-0.13
45	0.12	-0.17	-0.25	-0.11
46	0.03	0.11	0.20	0.18
47	0.27	0.06	-0.06	-0.16
48	-0.28	-0.18	-0.01	0.16
49	0.26	0.14	0.08	0.13
50	-0.10	-0.43	-0.45	-0.35

6. EEG Features – Normalization of EEG Asymmetry Ratio for Group 2

Subject	Asym_Delta	Asym_Theta	Asym_Alpha	Asym_Beta
1	0.12	0.08	0.06	0.12
2	0.04	0.05	0.03	0.08
3	0.10	0.08	-0.09	-0.23
4	0.20	0.16	0.18	0.52
5	-0.13	-0.10	-0.05	0.06
6	0.05	0.05	0.02	-0.03
7	0.06	0.02	-0.03	-0.13
8	-0.03	-0.07	-0.07	-0.08
9	0.05	0.04	-0.06	-0.17
10	0.14	0.07	0.05	0.01
11	0.04	0.00	-0.05	-0.25
12	0.05	0.04	-0.05	0.15
13	-0.08	0.04	0.04	0.01
14	-0.03	-0.07	-0.05	0.07
15	-0.04	-0.02	-0.05	-0.08
16	-0.06	-0.11	-0.07	0.02
17	0.03	0.04	0.01	-0.08
18	-0.08	-0.11	-0.09	0.03
19	-0.03	-0.05	-0.01	0.05
20	0.15	0.08	0.01	-0.07
21	-0.14	-0.15	-0.13	-0.01
22	0.02	-0.07	-0.04	0.27
23	0.04	0.06	0.08	-0.16
24	0.15	0.11	0.10	0.00
25	0.10	0.11	0.08	0.03
26	-0.11	0.03	0.08	0.24
27	0.02	0.01	0.01	-0.04
28	-0.11	-0.07	-0.03	-0.23
29	-0.04	-0.04	-0.02	0.07
30	-0.24	-0.15	-0.11	0.02

31	-0.10	-0.30	-0.37	-0.16
32	0.10	0.10	0.10	0.23
33	0.06	0.06	0.05	0.05
34	-0.05	-0.05	-0.03	-0.02
35	0.26	0.27	0.33	0.31
36	-0.20	-0.23	-0.24	-0.34
37	0.26	0.32	0.34	0.39
38	-0.17	-0.21	-0.16	-0.14
39	0.18	0.19	0.15	-0.40
40	0.05	0.07	0.08	0.08
41	0.01	0.14	0.28	0.30
42	0.09	0.06	0.09	0.02
43	-0.02	-0.02	0.00	0.05
44	-0.08	-0.04	-0.05	-0.43
45	-0.02	-0.03	-0.10	-0.33
46	0.20	0.13	0.15	0.38
47	0.37	0.28	0.25	0.12
48	0.20	0.16	0.18	0.49
49	-0.10	-0.06	-0.03	-0.22
50	0.17	0.08	0.07	0.29

7. EEG Features – Normalization of EEG Asymmetry Ratio for Group 3

Subject	Asym_Delta	Asym_Theta	Asym_Alpha	Asym_Beta
1	0.45	0.05	-0.05	-0.05
2	-0.29	-0.27	-0.16	-0.05
3	0.03	-0.01	0.09	0.06
4	0.21	0.14	-0.07	0.04
5	0.26	-0.06	-0.20	-0.34
6	-0.07	0.00	-0.01	-0.05
7	-0.25	-0.07	-0.08	-0.31
8	-0.03	-0.10	-0.03	0.03
9	0.11	0.07	0.01	-0.07
10	0.19	0.27	0.20	0.22
11	-0.37	-0.18	-0.03	0.02
12	-0.08	-0.06	-0.01	0.06
13	0.07	0.09	0.02	0.22
14	-0.03	-0.09	-0.03	-0.03
15	-0.15	-0.04	0.05	-0.05
16	-0.08	0.02	0.06	0.00
17	0.00	0.07	0.14	0.11
18	-0.16	-0.03	0.00	-0.14
19	-0.18	-0.08	-0.04	-0.03
20	0.03	0.02	-0.01	0.30
21	-0.09	-0.06	-0.06	-0.13
22	0.02	0.06	0.02	-0.02
23	-0.12	0.02	-0.31	-0.03
24	0.09	0.06	0.11	0.21
25	-0.08	-0.10	-0.10	-0.18
26	0.03	-0.02	-0.08	0.06
27	-0.16	-0.05	0.03	-0.03
28	-0.21	-0.28	-0.14	-0.37
29	-0.16	-0.03	0.00	-0.14
30	0.45	0.05	-0.05	-0.06
31	-0.29	-0.27	-0.16	-0.05
32	0.21	0.14	-0.07	0.05
33	0.05	-0.08	-0.09	-0.08
34	0.11	0.07	0.01	-0.07
35	0.20	0.26	0.19	0.21
36	-0.03	-0.09	-0.03	-0.03
37	0.00	0.07	0.14	0.11
38	-0.18	-0.08	-0.04	-0.03
39	0.02	0.06	0.02	-0.02
40	-0.16	-0.04	0.03	-0.02

8. EEG Features – Normalization of EEG Asymmetry Ratio for Group 4

Subject	Asym_Delta	Asym_Theta	Asym_Alpha	Asym_Beta
1	0.03	-0.21	-0.07	-0.16
2	0.20	0.06	-0.08	-0.01
3	-0.30	-0.29	-0.29	-0.20
4	-0.08	-0.06	-0.07	-0.02
5	0.06	0.06	0.08	-0.02
6	0.04	0.04	0.01	0.01
7	-0.10	-0.19	-0.14	-0.12
8	0.11	0.05	-0.01	0.03
9	-0.13	-0.10	-0.04	-0.22
10	0.16	0.07	0.03	0.02
11	0.05	0.05	0.05	-0.11
12	0.10	0.06	0.05	-0.04
13	0.12	-0.05	0.00	0.09
14	0.07	0.10	0.06	0.07
15	0.14	0.14	0.20	0.08
16	-0.16	-0.06	0.05	0.05
17	0.01	-0.02	-0.02	-0.08
18	0.09	0.09	0.15	0.14
19	0.03	0.01	0.02	-0.07
20	-0.17	-0.02	0.01	0.01
21	-0.11	-0.02	-0.02	0.19
22	0.19	0.09	0.02	-0.03
23	-0.03	-0.06	-0.11	-0.06
24	0.19	0.01	-0.01	0.05
25	0.09	0.07	0.13	0.17
26	-0.10	-0.06	-0.08	-0.17
27	0.08	0.08	0.07	0.23
28	-0.13	0.09	0.08	0.06
29	0.19	0.07	-0.06	-0.22
30	0.03	-0.21	-0.07	-0.15
31	0.20	0.06	-0.08	-0.01
32	0.06	0.07	0.08	-0.02
33	-0.10	-0.19	-0.13	-0.11
34	0.16	0.07	0.03	0.03
35	0.05	0.05	0.06	-0.12
36	0.14	0.14	0.20	0.73
37	0.08	0.08	0.15	0.14
38	-0.17	-0.02	0.01	0.01
39	-0.03	-0.06	-0.09	-0.05
40	0.08	0.08	0.07	0.22

9. **EEG Features – Relative Energy Ratio for Group**

Delta_log_ratio	**Theta_log_ratio**	**Alpha_log_ratio**	**Beta_log_ratio**
0.18	0.70	0.94	1.68
0.09	0.83	1.53	1.91
0.11	0.87	1.21	1.54
0.04	1.20	1.79	2.33
0.17	0.74	0.87	1.88
0.06	1.00	1.55	2.45
0.11	0.82	1.23	2.11
0.11	0.95	0.99	2.04
0.07	0.93	1.63	2.19
0.38	0.66	0.49	1.36
0.04	1.16	1.61	2.43
0.15	0.77	1.00	1.81
0.08	0.90	1.41	2.00
0.10	0.82	1.27	2.26
0.16	0.76	0.96	1.72
0.15	0.75	0.98	1.81
0.12	0.79	1.12	2.05
0.33	0.75	0.47	1.83
0.10	0.78	1.46	2.39
0.11	0.87	1.08	2.22
0.14	0.79	0.99	2.15
0.07	0.96	1.64	2.06
0.10	0.96	1.09	1.98
0.10	0.78	1.50	2.45
0.15	0.82	0.89	1.80
0.15	0.77	0.99	1.91
0.15	0.80	0.95	1.63
0.34	0.77	0.44	1.75
0.05	1.10	1.60	2.29
0.26	0.65	0.68	1.80
0.13	0.89	0.93	2.15
0.15	0.80	0.91	1.95
0.07	0.98	1.47	2.00
0.09	0.91	1.22	1.98
0.10	0.90	1.12	2.01
0.10	0.99	1.13	1.69
0.12	0.77	1.15	2.03
0.11	0.86	1.09	2.32
0.06	0.98	1.66	2.20
0.14	0.84	1.06	1.46
0.11	0.93	1.02	2.08

0.16	0.77	0.89	2.30
0.17	0.74	0.87	1.87
0.10	0.78	1.46	2.37
0.07	0.96	1.64	2.04
0.10	0.78	1.50	2.44
0.10	0.90	1.12	1.99
0.10	0.98	1.13	1.69
0.12	0.77	1.15	2.02
0.06	0.98	1.65	2.20

10. EEG Features – Relative Energy Ratio for Group 2

Delta_log_ratio	Theta_log_ratio	Alpha_log_ratio	Beta_log_ratio
0.11	0.82	1.34	1.53
0.12	0.75	1.34	2.00
0.11	0.80	1.38	1.81
0.12	0.77	1.44	1.53
0.12	0.74	1.42	1.67
0.11	0.74	1.47	2.05
0.11	0.76	1.45	2.01
0.10	0.81	1.44	1.89
0.14	0.74	1.23	1.51
0.13	0.75	1.36	1.51
0.09	0.86	1.45	1.66
0.12	0.76	1.19	1.94
0.10	0.80	1.44	2.05
0.13	0.68	1.38	2.00
0.11	0.73	1.40	2.24
0.13	0.67	1.35	2.20
0.11	0.77	1.41	1.96
0.11	0.73	1.47	2.19
0.08	0.85	1.55	2.16
0.11	0.77	1.43	2.05
0.10	0.79	1.49	2.13
0.07	0.93	1.57	1.98
0.11	0.79	1.43	1.79
0.10	0.80	1.48	1.96
0.10	0.79	1.51	2.09
0.08	0.87	1.70	1.96
0.09	0.84	1.53	2.14
0.11	0.78	1.44	1.87
0.10	0.79	1.55	2.17
0.09	0.83	1.54	2.12

0.15	0.68	1.23	1.52
0.13	0.69	1.37	1.91
0.09	0.82	1.47	2.09
0.09	0.80	1.60	2.21
0.10	0.78	1.49	2.04
0.10	0.78	1.40	2.15
0.11	0.77	1.43	1.90
0.14	0.66	1.26	2.11
0.12	0.76	1.41	1.75
0.09	0.80	1.52	2.21
0.12	0.73	1.39	2.19
0.09	0.82	1.55	2.26
0.12	0.73	1.40	1.87
0.11	0.78	1.44	1.76
0.14	0.80	1.22	1.19
0.10	0.81	1.45	1.91
0.10	0.78	1.52	2.29
0.12	0.77	1.45	1.53
0.11	0.78	1.43	1.87
0.10	0.79	1.43	1.91

11. EEG Features – Relative Energy Ratio for Group 3

Delta_log_ratio	Theta_log_ratio	Alpha_log_ratio	Beta_log_ratio
0.07	1.06	1.29	1.92
0.07	0.92	1.56	2.06
0.08	0.90	1.42	2.04
0.10	0.85	1.30	1.93
0.29	0.72	0.62	1.27
0.14	0.79	1.05	1.68
0.29	0.59	0.67	1.64
0.10	0.88	1.29	1.84
0.07	0.92	1.53	2.28
0.09	0.83	1.43	2.29
0.06	0.98	1.80	2.51
0.08	0.90	1.42	2.38
0.09	1.01	1.11	2.12
0.12	0.74	1.28	2.25
0.25	0.62	0.73	1.83
0.10	0.84	1.32	1.88
0.21	0.76	0.75	1.46
0.14	0.81	0.99	1.87
0.15	0.77	0.98	1.90
0.15	0.82	0.87	2.19
0.05	1.06	1.89	2.50
0.08	0.88	1.52	2.04
0.10	0.83	1.36	2.07
0.09	0.87	1.40	2.23
0.11	0.84	1.15	1.94
0.08	0.85	1.79	2.38
0.09	0.86	1.37	2.32
0.18	0.82	0.79	1.70
0.14	0.81	0.99	1.87
0.07	1.06	1.28	1.91
0.07	0.92	1.56	2.05
0.10	0.85	1.29	1.92
0.10	0.86	1.18	2.12
0.07	0.91	1.53	2.26
0.09	0.83	1.43	2.28
0.12	0.74	1.28	2.24
0.21	0.75	0.75	1.46
0.15	0.77	0.97	1.89
0.08	0.88	1.51	2.01
0.09	0.86	1.36	2.29

12. EEG Features – Relative Energy Ratio for Group 4

Delta_log_ratio	Theta_log_ratio	Alpha_log_ratio	Beta_log_ratio
0.17	0.78	0.87	1.55
0.08	0.90	1.49	2.04
0.09	0.83	1.47	2.02
0.18	0.61	1.12	1.84
0.10	0.77	1.42	2.34
0.12	0.85	1.09	1.70
0.07	0.89	1.64	2.53
0.21	0.63	0.84	2.09
0.10	0.92	1.12	1.81
0.09	0.85	1.39	2.01
0.10	0.81	1.35	2.20
0.06	0.93	1.81	2.51
0.12	0.85	1.08	1.93
0.11	0.81	1.26	1.89
0.17	0.75	0.89	1.96
0.22	0.72	0.71	1.95
0.15	0.75	1.04	1.69
0.24	0.73	0.72	1.40
0.30	0.65	0.60	1.55
0.22	0.62	0.82	1.99
0.16	0.79	0.86	2.23
0.10	0.81	1.37	1.86
0.09	0.88	1.43	1.84
0.08	0.89	1.43	1.97
0.10	0.77	1.42	2.38
0.14	0.75	1.12	1.75
0.10	0.74	1.73	2.40
0.13	0.73	1.24	1.92
0.32	0.78	0.48	1.56
0.17	0.78	0.87	1.54
0.08	0.90	1.48	2.03
0.10	0.77	1.42	2.33
0.07	0.89	1.64	2.52
0.09	0.85	1.38	1.99
0.10	0.81	1.34	2.19
0.16	0.75	0.89	2.16
0.24	0.73	0.72	1.39
0.22	0.62	0.82	1.98
0.09	0.87	1.42	1.82
0.10	0.74	1.73	2.38

13. EEG Features – Entropy for Group 1, 2, 3, 4

Entropy G1	Entropy G2	Entropy G3	Entropy G4
0.40	0.32	0.24	0.41
0.26	0.30	0.23	0.25
0.32	0.30	0.26	0.27
0.15	0.32	0.29	0.38
0.39	0.31	0.50	0.28
0.21	0.29	0.36	0.33
0.30	0.29	0.48	0.22
0.31	0.28	0.29	0.42
0.22	0.35	0.23	0.31
0.52	0.33	0.26	0.27
0.17	0.28	0.19	0.28
0.36	0.33	0.24	0.20
0.26	0.28	0.27	0.32
0.29	0.32	0.30	0.30
0.38	0.29	0.45	0.38
0.37	0.32	0.29	0.43
0.32	0.29	0.45	0.37
0.47	0.29	0.35	0.47
0.27	0.25	0.36	0.49
0.31	0.29	0.36	0.43
0.34	0.27	0.17	0.37
0.22	0.23	0.25	0.29
0.29	0.29	0.28	0.27
0.26	0.28	0.26	0.26
0.37	0.27	0.31	0.28
0.36	0.24	0.22	0.35
0.38	0.25	0.26	0.25
0.48	0.29	0.40	0.33
0.18	0.26	0.35	0.48
0.46	0.25	0.24	0.41
0.33	0.37	0.23	0.25
0.37	0.32	0.29	0.28
0.23	0.27	0.29	0.22
0.28	0.25	0.23	0.27
0.30	0.28	0.26	0.28
0.30	0.28	0.30	0.38
0.32	0.29	0.45	0.47
0.30	0.33	0.36	0.43

0.21	0.31	0.25	0.27
0.36	0.26	0.26	0.26
0.31	0.30		
0.37	0.25		
0.39	0.31		
0.27	0.30		
0.22	0.38		
0.26	0.28		
0.30	0.27		
0.30	0.32		
0.32	0.29		
0.21	0.29		

14. FKM Centroid for Group 1

Delta	Theta	Alpha	Beta
0.17317	0.74126	0.87225	1.87985
0.10753	0.81557	1.23206	2.10705
0.10925	0.95129	0.99424	2.03605
0.14595	0.76906	1.00153	1.80719
0.15812	0.75525	0.95772	1.71542
0.15264	0.75441	0.97983	1.80817
0.12349	0.78721	1.12258	2.05287
0.11025	0.87162	1.07686	2.21945
0.13759	0.79076	0.98919	2.14822
0.09747	0.95839	1.09414	1.9806
0.15141	0.82153	0.89438	1.79591
0.14514	0.82153	0.89438	1.79591
0.12728	0.88767	0.93007	2.15234
0.15075	0.79637	0.91274	1.95071
0.10271	0.90425	1.11809	2.01127
0.09530	0.98583	1.13374	1.69417
0.12418	0.77467	1.14717	2.02553
0.11008	0.86039	1.09025	2.32365
0.10924	0.93001	1.01538	2.0761
0.1569	0.7685	0.89343	2.30236
0.17419	0.74052	0.86902	1.87123
0.10308	0.90446	1.11605	1.99448
0.09576	0.9838	1.13236	1.69131
0.12442	0.77422	1.14698	2.01621

15. FKM centroids for Group 2

Delta	Theta	Alpha	Beta
0.17752	0.70132	0.93646	1.68107
0.09055	0.83433	1.52972	1.9131
0.11146	0.86798	1.20642	1.5426
0.03793	1.20161	1.79211	2.3343
0.17317	0.74126	0.87225	1.87985
0.06118	1.00172	1.54966	2.44512
0.10753	0.81557	1.23206	2.10705
0.10925	0.95129	0.99424	2.03605
0.0697	0.92637	1.63076	2.19415
0.38259	0.65963	0.49106	1.35809
0.04454	1.15976	1.61089	2.42578
0.08317	0.90068	1.41379	1.99902
0.10297	0.81796	1.27129	2.26236
0.12349	0.78721	1.12258	2.05287
0.11025	0.87162	1.07686	2.21945
0.06659	0.95556	1.64394	2.06141
0.09747	0.95839	1.09414	1.9806
0.09819	0.77645	1.50177	2.45048
0.14514	0.77173	0.98927	1.90522
0.34458	0.76825	0.44423	1.75362
0.12728	0.88767	0.93007	2.15234
0.15075	0.79637	0.91274	1.95071
0.10271	0.90425	1.11809	2.01127
0.12418	0.77467	1.14717	2.02553
0.11008	0.86039	1.09025	2.32365
0.0618	0.98168	1.65616	2.20394
0.17419	0.74052	0.86902	1.87123
0.09883	0.78353	1.46011	2.37125
0.09826	0.77647	1.50097	2.44095
0.09576	0.9838	1.13236	1.69131
0.12442	0.77422	1.14698	2.01621
0.06216	0.98001	1.65232	2.19522

16. FKM centroids for Group 3

Delta	Theta	Alpha	Beta
0.17317	0.74126	0.87225	1.87985
0.10753	0.81557	1.23206	2.10705
0.04454	1.15976	1.61089	2.42578
0.14595	0.76906	1.00153	1.80719
0.15141	0.82153	0.89438	1.79591
0.15083	0.80163	0.95071	1.62921
0.15075	0.79637	0.91274	1.95071
0.0698	0.98081	1.46977	1.99778
0.10271	0.90425	1.11809	2.01127
0.13505	0.83552	1.06202	1.46195

17. FKM centroids for Group 4

Delta	Theta	Alpha	Beta
0.17317	0.74126	0.87225	1.87985
0.10753	0.81557	1.23206	2.10705
0.04454	1.15976	1.61089	2.42578
0.14595	0.76906	1.00153	1.80719
0.15141	0.82153	0.89438	1.79591
0.15083	0.80163	0.95071	1.62921
0.15075	0.79637	0.91274	1.95071
0.0698	0.98081	1.46977	1.99778
0.10271	0.90425	1.11809	2.01127
0.13505	0.83552	1.06202	1.46195

APPENDIX D

Equipment Validation Data

EXP NO.	CHANNEL	FFMSG	FRBG	∆ FREQ.	∆ FREQ.CH1-CH2	MARK
1	CH1	0.5	0.5	0	0	CH1 = CH2
	CH2	0.5	0.5	0		
2	CH1	1	1	0	0	CH1 = CH2
	CH2	1	1	0		
3	CH1	2	2	0	0	CH1 = CH2
	CH2	2	2	0		
4	CH1	5	5	0	0	CH1 = CH2
	CH2	5	5	0		
5	CH1	10	10	0	0	CH1 = CH2
	CH2	10	10	0		
6	CH1	15	15	0	0	CH1 = CH2
	CH2	15	15	0		
7	CH1	20	20	0	0	CH1 = CH2
	CH2	20	20	0		
8	CH1	30	30	0	0	CH1 = CH2
	CH2	30	30	0		
9	CH1	40	40	0	0	CH1 = CH2
	CH2	40	40	0		
10	CH1	50	50	0	0	CH1 = CH2
	CH2	50	50	0		
11	CH1	0.5	0.5	0	0	CH1 = CH2
	CH2	0.5	0.5	0		
12	CH1	1	1	0	0	CH1 = CH2
	CH2	1	1	0		
13	CH1	2	2	0	0	CH1 = CH2
	CH2	2	2	0		
14	CH1	5	5	0	0	CH1 = CH2
	CH2	5	5	0		
15	CH1	10	10	0	0	CH1 = CH2
	CH2	10	10	0		
16	CH1	15	15	0	0	CH1 = CH2
	CH2	15	15	0		
17	CH1	20	20	0	0	CH1 = CH2
	CH2	20	20	0		
18	CH1	30	30	0	0	CH1 = CH2
	CH2	30	30	0		
19	CH1	40	40	0	0	CH1 = CH2
	CH2	40	40	0		
20	CH1	50	50	0	0	CH1 = CH2
	CH2	50	50	0		

21	**CH1**	0.5	0.5	0	0	**CH1 = CH2**
	CH2	0.5	0.5	0		
22	**CH1**	1	1	0	0	**CH1 = CH2**
	CH2	1	1	0		
23	**CH1**	2	2	0	0	**CH1 = CH2**
	CH2	2	2	0		
24	**CH1**	5	5	0	0	**CH1 = CH2**
	CH2	5	5	0		
25	**CH1**	10	10	0	0	**CH1 = CH2**
	CH2	10	10	0		
26	**CH1**	15	15	0	0	**CH1 = CH2**
	CH2	15	15	0		
27	**CH1**	20	20	0	0	**CH1 = CH2**
	CH2	20	20	0		
28	**CH1**	30	30	0	0	**CH1 = CH2**
	CH2	30	30	0		
29	**CH1**	40	40	0	0	**CH1 = CH2**
	CH2	40	40	0		
30	**CH1**	50	50	0	0	**CH1 = CH2**
	CH2	50	50	0		
31	**CH1**	0.5	0.5	0	0	**CH1 = CH2**
	CH2	0.5	0.5	0		
32	**CH1**	1	1	0	0	**CH1 = CH2**
	CH2	1	1	0		
33	**CH1**	2	2	0	0	**CH1 = CH2**
	CH2	2	2	0		
34	**CH1**	5	5	0	0	**CH1 = CH2**
	CH2	5	5	0		
35	**CH1**	10	10	0	0	**CH1 = CH2**
	CH2	10	10	0		
36	**CH1**	15	15	0	0	**CH1 = CH2**
	CH2	15	15	0		
37	**CH1**	20	20	0	0	**CH1 = CH2**
	CH2	20	20	0		
38	**CH1**	30	30	0	0	**CH1 = CH2**
	CH2	30	30	0		
39	**CH1**	40	40	0	0	**CH1 = CH2**
	CH2	40	40	0		
40	**CH1**	50	50	0	0	**CH1 = CH2**
	CH2	50	50	0		
41	**CH1**	0.5	0.5	0	0	**CH1 = CH2**
	CH2	0.5	0.5	0		
42	**CH1**	1	1	0	0	**CH1 = CH2**
	CH2	1	1	0		
43	**CH1**	2	2	0	0	**CH1 =**

	CH2	2	2	0		CH2
44	CH1	5	5	0	0	CH1 = CH2
	CH2	5	5	0		
45	CH1	10	10	0	0	CH1 = CH2
	CH2	10	10	0		
46	CH1	15	15	0	0	CH1 = CH2
	CH2	15	15	0		
47	CH1	20	20	0	0	CH1 = CH2
	CH2	20	20	0		
48	CH1	30	30	0	0	CH1 = CH2
	CH2	30	30	0		
49	CH1	40	40	0	0	CH1 = CH2
	CH2	40	40	0		
50	CH1	50	50	0	0	CH1 = CH2
	CH2	50	50	0		
51	CH1	0.5	0.5	0	0	CH1 = CH2
	CH2	0.5	0.5	0		
52	CH1	1	1	0	0	CH1 = CH2
	CH2	1	1	0		
53	CH1	2	2	0	0	CH1 = CH2
	CH2	2	2	0		
54	CH1	5	5	0	0	CH1 = CH2
	CH2	5	5	0		
55	CH1	10	10	0	0	CH1 = CH2
	CH2	10	10	0		
56	CH1	15	15	0	0	CH1 = CH2
	CH2	15	15	0		
57	CH1	20	20	0	0	CH1 = CH2
	CH2	20	20	0		
58	CH1	30	30	0	0	CH1 = CH2
	CH2	30	30	0		
59	CH1	40	40	0	0	CH1 = CH2
	CH2	40	40	0		
60	CH1	50	50	0	0	CH1 = CH2
	CH2	50	50	0		
61	CH1	0.5	0.5	0	0	CH1 = CH2
	CH2	0.5	0.5	0		
62	CH1	1	1	0	0	CH1 = CH2
	CH2	1	1	0		
63	CH1	2	2	0	0	CH1 = CH2
	CH2	2	2	0		
64	CH1	5	5	0	0	CH1 = CH2
	CH2	5	5	0		
65	CH1	10	10	0	0	CH1 = CH2
	CH2	10	10	0		

66	**CH1**	15	15	0	0	**CH1 = CH2**
	CH2	15	15	0		
67	**CH1**	20	20	0	0	**CH1 = CH2**
	CH2	20	20	0		
68	**CH1**	30	30	0	0	**CH1 = CH2**
	CH2	30	30	0		
69	**CH1**	40	40	0	0	**CH1 = CH2**
	CH2	40	40	0		
70	**CH1**	50	50	0	0	**CH1 = CH2**
	CH2	50	50	0		
71	**CH1**	0.5	0.5	0	0	**CH1 = CH2**
	CH2	0.5	0.5	0		
72	**CH1**	1	1	0	0	**CH1 = CH2**
	CH2	1	1	0		
73	**CH1**	2	2	0	0	**CH1 = CH2**
	CH2	2	2	0		
74	**CH1**	5	5	0	0	**CH1 = CH2**
	CH2	5	5	0		
75	**CH1**	10	10	0	0	**CH1 = CH2**
	CH2	10	10	0		
76	**CH1**	15	15	0	0	**CH1 = CH2**
	CH2	15	15	0		
77	**CH1**	20	20	0	0	**CH1 = CH2**
	CH2	20	20	0		
78	**CH1**	30	30	0	0	**CH1 = CH2**
	CH2	30	30	0		
79	**CH1**	40	40	0	0	**CH1 = CH2**
	CH2	40	40	0		
80	**CH1**	50	50	0	0	**CH1 = CH2**
	CH2	50	50	0		
81	**CH1**	0.5	0.5	0	0	**CH1 = CH2**
	CH2	0.5	0.5	0		
82	**CH1**	1	1	0	0	**CH1 = CH2**
	CH2	1	1	0		
83	**CH1**	2	2	0	0	**CH1 = CH2**
	CH2	2	2	0		
84	**CH1**	5	5	0	0	**CH1 = CH2**
	CH2	5	5	0		
85	**CH1**	10	10	0	0	**CH1 = CH2**
	CH2	10	10	0		
86	**CH1**	15	15	0	0	**CH1 = CH2**
	CH2	15	15	0		
87	**CH1**	20	20	0	0	**CH1 = CH2**
	CH2	20	20	0		
88	**CH1**	30	30	0	0	**CH1 =**

	CH2	30	30	0		CH2
89	CH1	40	40	0	0	CH1 = CH2
	CH2	40	40	0		
90	CH1	50	50	0	0	CH1 = CH2
	CH2	50	50	0		

VFMSG	VRBG	∆ VOLTAGE (VFMSG - VRBG)	∆ VOLTAGE CH1-CH2	MARK
100	198.906534	-98.90653396	2.638168424	CH2 > CH1
100	201.544702	-101.5447024		
100	265.323312	-165.3233118	-0.154768617	CH1 > CH2
100	265.168543	-165.1685432		
100	283.489372	-183.4893721	-0.274650092	CH1 > CH2
100	283.214722	-183.214722		
100	246.251004	-146.2510039	0	CH1 = CH2
100	246.251004	-146.2510039		
100	126.890168	-26.89016837	0	CH1 = CH2
100	126.890168	-26.89016837		
100	163.7941	-63.79409985	-13.22124018	CH1 > CH2
100	150.57286	-50.57285967		
100	285.300032	-185.3000318	-2.535715479	CH1 > CH2
100	282.764316	-182.7643163		
100	299.309365	-199.309365	3.844535039	CH2 > CH1
100	303.1539	-203.1539		
100	47.4031801	52.59681991	0.62479218	CH2 > CH1
100	48.0279723	51.97202773		
100	9.28244074	90.71755926	0.327738529	CH2 > CH1
100	9.61017926	90.38982074		
150	294.216481	-144.2164813	3.020099658	CH2 > CH1
150	297.236581	-147.236581		
150	393.767686	-243.7676858	0.626497106	CH2 > CH1
150	394.394183	-244.3941829		
150	423.180057	-273.1800569	-0.398252139	CH1 > CH2
150	422.781805	-272.7818047		
150	360.40609	-210.4060897	1.058669163	CH2 > CH1
150	361.464759	-211.4647589		
150	170.38903	-20.38902992	11.86735474	CH2 > CH1
150	182.256385	-32.25638466		
150	250.987359	-100.9873592	3.036792185	CH2 > CH1
150	254.024151	-104.0241514		
150	418.968224	-268.9682238	-1.642473863	CH1 > CH2
150	417.32575	-267.3257499		
150	440.586959	-290.5869585	-1.844210801	CH1 > CH2
150	438.742748	-288.7427477		
150	69.5431167	80.45688331	-0.587936068	CH1 > CH2
150	68.9551806	81.04481938		
150	13.4553639	136.5446361	-2.840538174	CH1 >

150	10.6148258	139.3851742		CH2
200	391.895256	-191.8952557	3.050726345	CH2 >
200	394.945982	-194.9459821		CH1
200	522.908756	-322.9087563	-0.004177644	CH1 >
200	522.904579	-322.9045786		CH2
200	583.481111	-383.4811105	-1.968108689	CH1 >
200	581.513002	-381.5130018		CH2
200	475.929526	-275.9295258	-3.435928053	CH1 >
200	472.493598	-272.4935977		CH2
200	238.364259	-38.36425917	4.774144192	CH2 >
200	243.138403	-43.13840336		CH1
200	282.529862	-82.52986155	-8.522136624	CH1 >
200	274.007725	-74.00772493		CH2
200	566.952748	-366.9527484	-4.836322644	CH1 >
200	562.116426	-362.1164258		CH2
200	595.993228	-395.9932279	-4.159657765	CH1 >
200	591.83357	-391.8335702		CH2
200	87.9909739	112.0090261	0.619225406	CH2 >
200	88.6101993	111.3898007		CH1
200	10.4204456	189.5795544	1.128111612	CH2 >
200	11.5485572	188.4514428		CH1
250	515.529533	-265.5295327	-2.752523326	CH1 >
250	512.777009	-262.7770093		CH2
250	681.610959	-431.6109592	1.553131977	CH2 >
250	683.164091	-433.1640912		CH1
250	705.770713	-455.7707131	-3.346337098	CH1 >
250	702.424376	-452.424376		CH2
250	602.910046	-352.9100457	-3.412992305	CH1 >
250	599.497053	-349.4970534		CH2
250	307.055378	-57.05537803	0.626092785	CH2 >
250	307.681471	-57.68147082		CH1
250	364.530121	-114.5301211	5.611493088	CH2 >
250	370.141614	-120.1416141		CH1
250	701.734246	-451.734246	-1.230310866	CH1 >
250	700.503935	-450.5039351		CH2
250	745.495567	-495.4955673	-0.952757325	CH1 >
250	744.54281	-494.54281		CH2
250	110.146544	139.853456	-1.295525466	CH1 >
250	108.851019	141.1489815		CH2
250	14.3176879	235.6823121	0.475138198	CH2 >
250	14.7928261	235.2071739		CH1
300	590.073171	-290.0731706	4.612364068	CH2 >
300	594.685535	-294.6855346		CH1
300	802.912935	-502.9129354	-1.838218842	CH1 >
300	801.074717	-501.0747165		CH2
300	851.029502	-551.0295022	-2.029040834	CH1 >
300	849.000461	-549.0004614		CH2
300	725.422402	-425.4224021	-0.447734858	CH1 >
300	724.974667	-424.9746673		CH2

300	361.583297	-61.58329702	-15.80553829	CH1 > CH2
300	345.777759	-45.77775872		
300	489.522046	-189.5220458	-16.45922267	CH1 > CH2
300	473.062823	-173.0628231		
300	859.448801	-559.4488015	-1.199629463	CH1 > CH2
300	858.249172	-558.249172		
300	881.539557	-581.5395567	-2.372455963	CH1 > CH2
300	879.167101	-579.1671007		
300	131.078197	168.9218029	1.113983295	CH2 > CH1
300	132.19218	167.8078196		
300	18.0196194	281.9803806	0.447067672	CH2 > CH1
300	18.4666871	281.5333129		
350	686.145607	-336.1456071	2.25506016	CH2 > CH1
350	688.400667	-338.4006673		
350	955.629656	-605.6296564	-2.049797474	CH1 > CH2
350	953.579859	-603.5798589		
350	986.655579	-636.6555794	-2.138612432	CH1 > CH2
350	984.516967	-634.516967		
350	833.44533	-483.4453305	-0.918058769	CH1 > CH2
350	832.527272	-482.5272717		
350	483.979756	-133.9797558	6.8100501	CH2 > CH1
350	490.789806	-140.7898059		
350	548.581566	-198.5815664	12.45452783	CH2 > CH1
350	561.036094	-211.0360942		
350	1006.35922	-656.3592246	-7.1568682	CH1 > CH2
350	999.202356	-649.2023564		
350	1023.49984	-673.4998378	0.791467773	CH2 > CH1
350	1024.29131	-674.2913055		
350	150.304901	199.6950987	-0.747394519	CH1 > CH2
350	149.557507	200.4424933		
350	20.5984064	329.4015936	-1.210917402	CH1 > CH2
350	19.387489	330.612511		
400	783.183202	-383.1832024	6.63435015	CH2 > CH1
400	789.817553	-389.8175525		
400	1049.31038	-649.3103848	-1.516767776	CH1 > CH2
400	1047.79362	-647.793617		
400	1122.01437	-722.014372	-2.939744377	CH1 > CH2
400	1119.07463	-719.0746277		
400	958.168976	-558.1689756	1.503730422	CH2 > CH1
400	959.672706	-559.672706		
400	486.757147	-86.75714712	-13.69533231	CH1 > CH2
400	473.061815	-73.0618148		
400	581.874742	-181.8747422	-8.939811051	CH1 > CH2
400	572.934931	-172.9349312		
400	1147.08052	-747.0805233	-12.75264745	CH1 > CH2
400	1134.32788	-734.3278758		
400	1170.92098	-770.9209803	1.418259229	CH2 > CH1
400	1172.33924	-772.3392396		
400	175.775258	224.2247423	-0.934363297	CH1 >

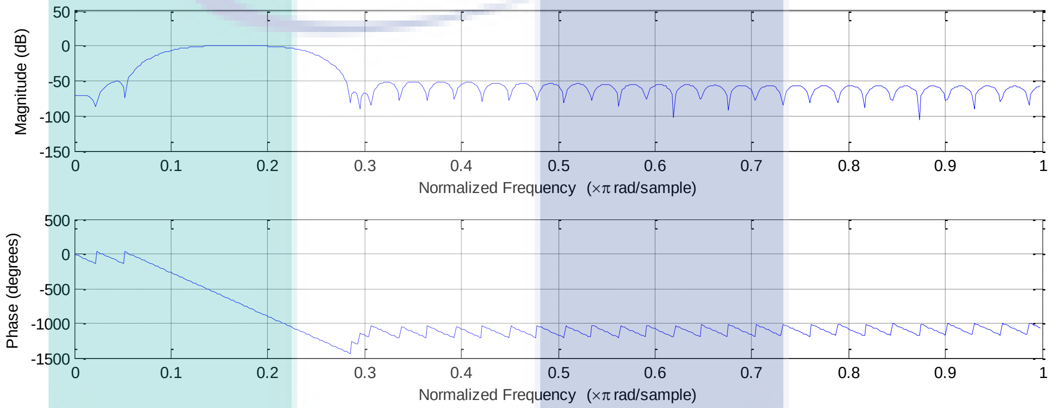
400	174.840894	225.1591056		CH2
400	17.5212633	382.4787367	1.029132796	CH2 > CH1
400	18.5503961	381.4496039		
450	882.716782	-432.7167819	5.558824775	CH2 > CH1
450	888.275607	-438.2756066		
450	1191.06589	-741.0658914	3.010374488	CH2 > CH1
450	1194.07627	-744.0762659		
450	1256.36325	-806.3632496	0	CH1 = CH2
450	1256.36325	-806.3632496		
450	1103.11493	-653.114926	-11.56103359	CH1 > CH2
450	1091.55389	-641.5538924		
450	528.044199	-78.04419886	17.07116805	CH2 > CH1
450	545.115367	-95.11536691		
450	696.298635	-246.2986349	-7.488616173	CH1 > CH2
450	688.810019	-238.8100187		
450	1251.93173	-801.9317324	-3.536797423	CH1 > CH2
450	1248.39494	-798.394935		
450	1332.64329	-882.6432895	-10.83191826	CH1 > CH2
450	1321.81137	-871.8113713		
450	192.779788	257.2202118	2.001130588	CH2 > CH1
450	194.780919	255.2190812		
450	28.175638	421.824362	-0.901518087	CH1 > CH2
450	27.2741199	422.7258801		
500	974.378882	-474.3788825	5.750319525	CH2 > CH1
500	980.129202	-480.129202		
500	1293.73281	-793.7328108	-9.664294788	CH1 > CH2
500	1284.06852	-784.068516		
500	1375.74051	-875.7405109	-10.6195506	CH1 > CH2
500	1365.12096	-865.1209603		
500	1229.22895	-729.2289508	-10.60538846	CH1 > CH2
500	1218.62356	-718.6235623		
500	672.37958	-172.3795799	20.99522872	CH2 > CH1
500	693.374809	-193.3748087		
500	720.741109	-220.7411087	3.56548009	CH2 > CH1
500	724.306589	-224.3065888		
500	1378.56231	-878.5623106	-10.37927293	CH1 > CH2
500	1368.18304	-868.1830377		
500	1426.71313	-926.7131268	-2.11428105	CH1 > CH2
500	1424.59885	-924.5988457		
500	231.231724	268.7682764	-0.906658537	CH1 > CH2
500	230.325065	269.674935		
500	25.7814435	474.2185565	4.886433285	CH2 > CH1
500	30.6678768	469.3321232		
		AVERAGE VOLTAGE DIFFERENCE	**-0.959848727**	

APPENDIX E

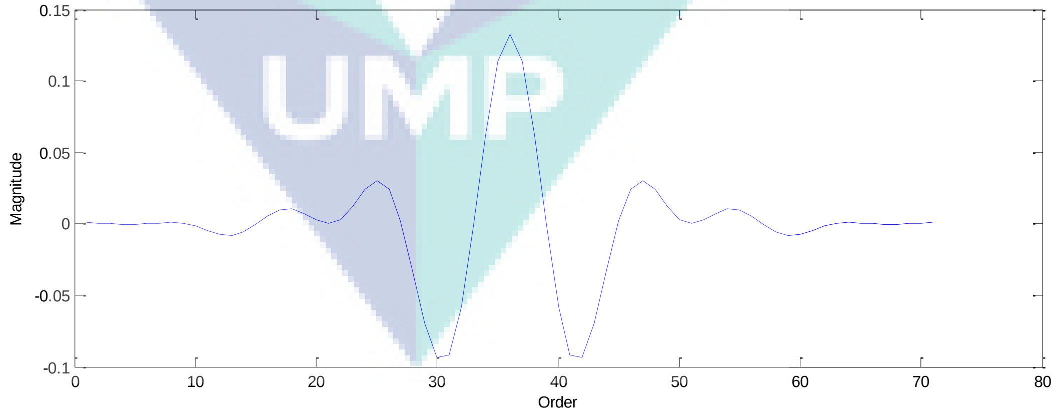
The process to design and implement Window-based Filter to Capture Beta band from EEG Signals

Type of EEG Band = Beta band [13 – 30 Hz]
Sampling Frequency=256 Hz,
Maximum Frequency =256/2=128 Hz
Frequency Cut-Off= [0.1016 0.2344]
Type of Filter = Bandpass
Window=Hamming

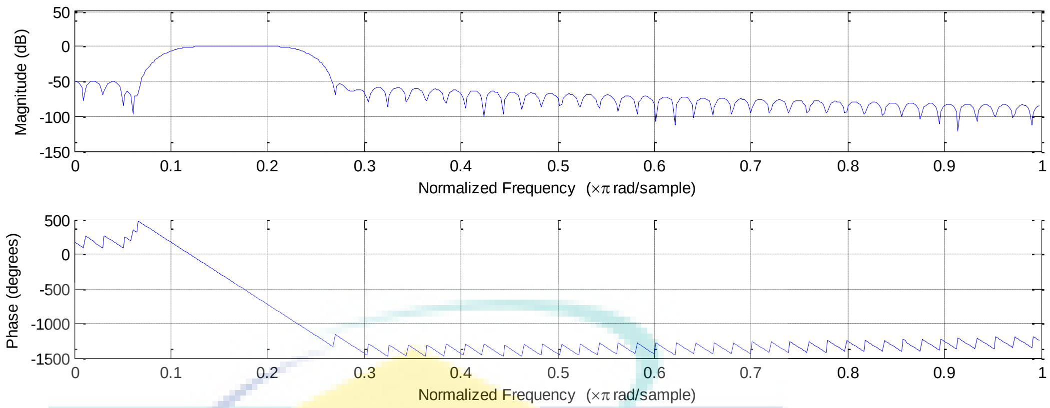
Plot of Frequency Response with Filter Order, N = 70



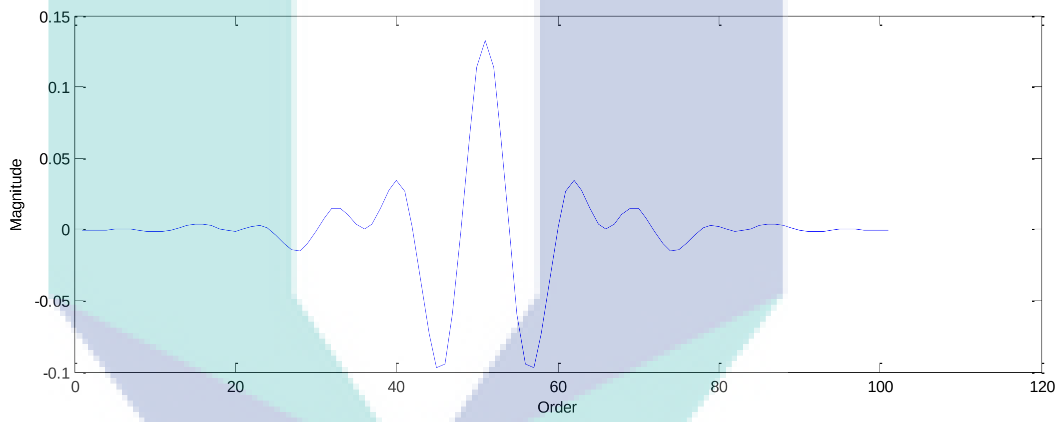
Plot of Filter Coefficient



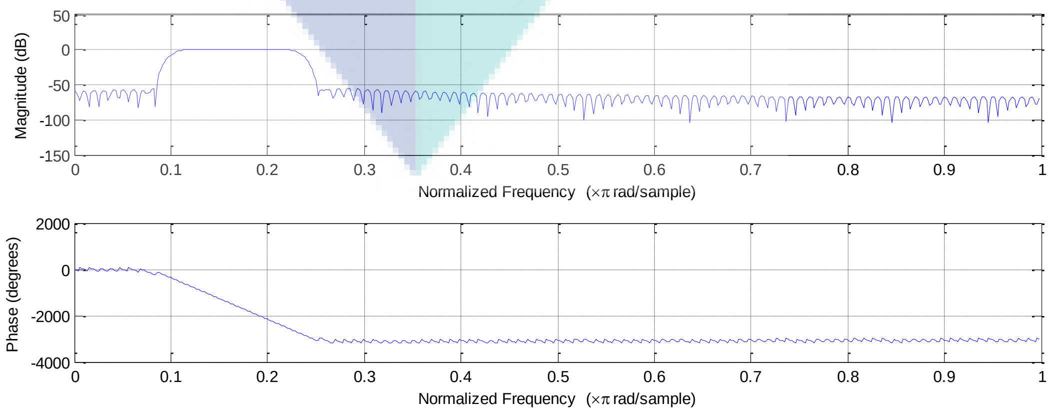
Plot of Frequency Response with Filter Order, N = 100



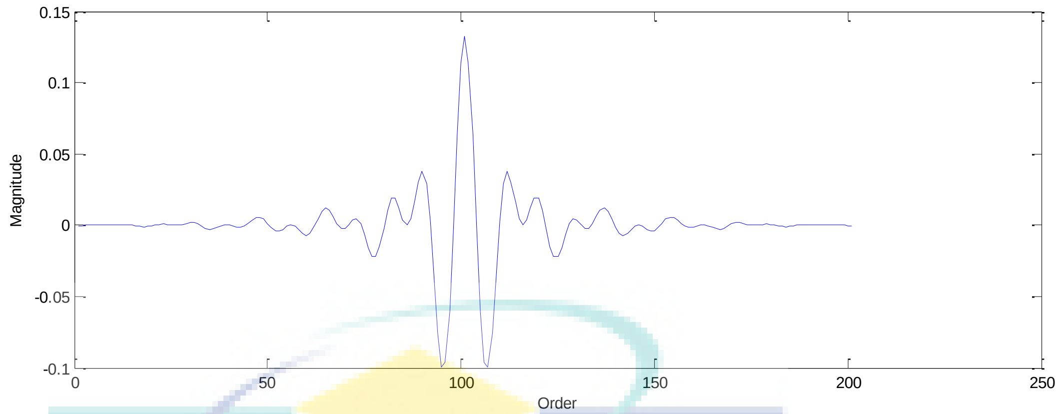
Plot of Filter Coefficient



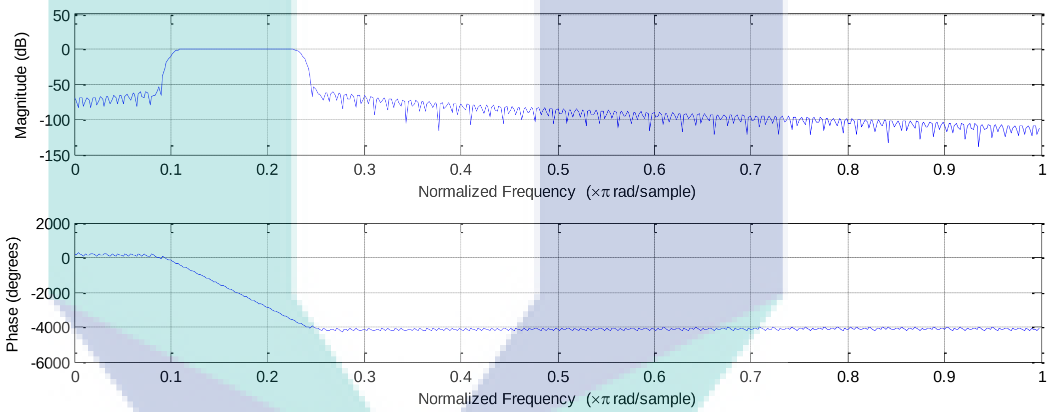
Plot of Frequency Response with Filter Order, N = 200



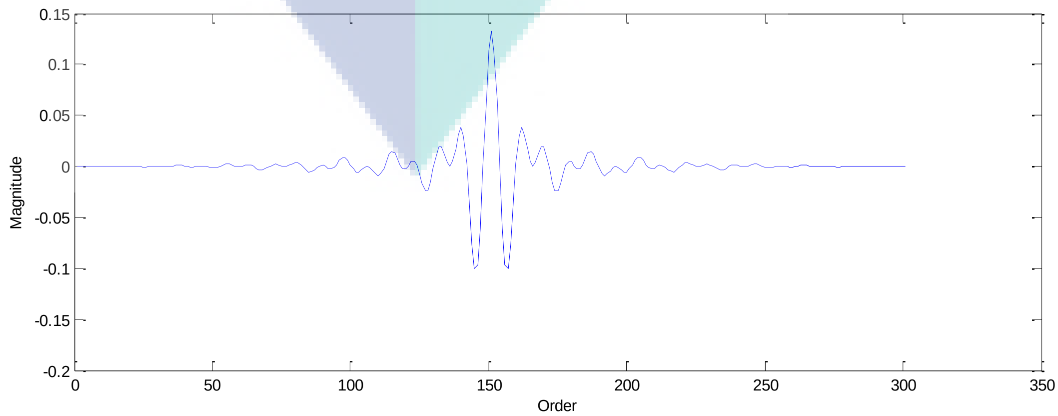
Plot of Filter Coefficient



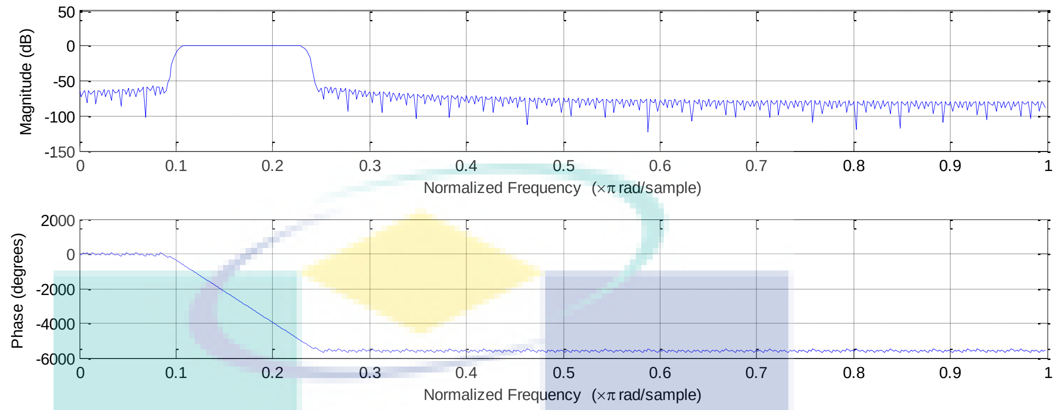
Plot of Frequency Response with Filter Order, N = 300



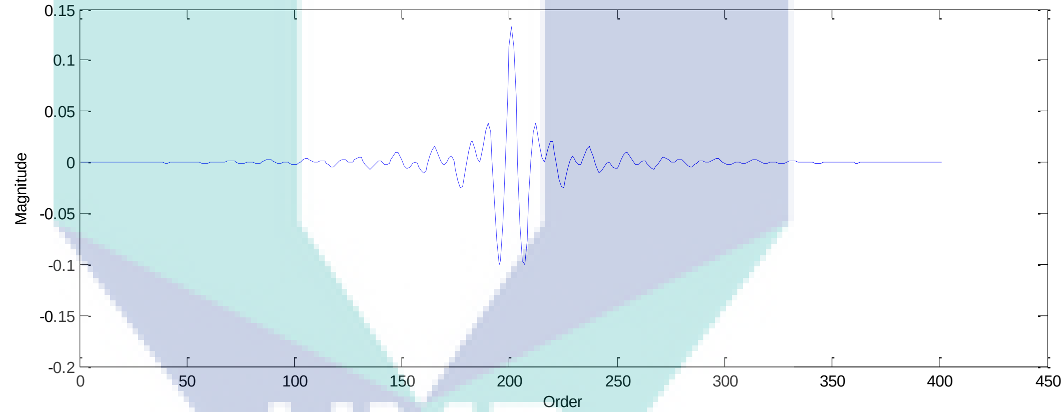
Plot of Filter Coefficient



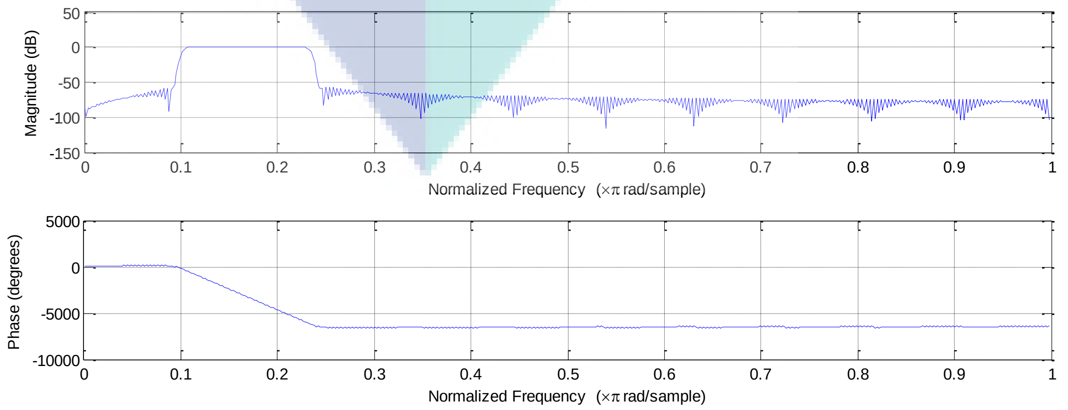
Plot of Frequency Response with Filter Order, N = 400



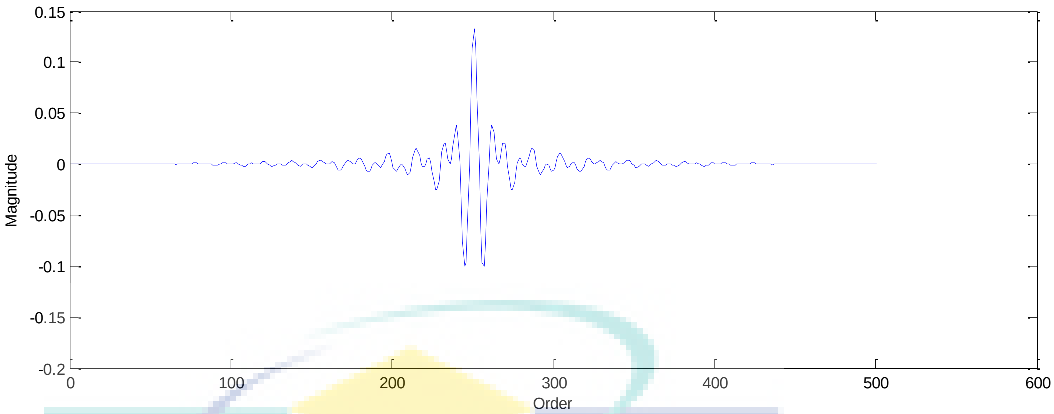
Plot of Filter Coefficient



Plot of Frequency Response with Filter Order, N = 500



Plot of Filter Coefficient



It can be oberserved that there are difference of Frequency responses and filter coefficients for each window when N is set to 500.

APPENDIX F

1. MATLAB Coding to Analyze and Filter EEG Signals

```
clear all;
load('rightch1.mat');
rightbeforefilterx=rightbeforefilter';
load('leftch2.mat');
leftbeforefilterx=leftbeforefilter';

% Reset the Offset of raw data
A_R=rightbeforefilterx(1:end,2);
mean_rhs=mean(A_R);
B_L=leftbeforefilterx(1:end,2);
mean_lhs=mean(B_L);
if abs(mean_rhs) > abs(mean_lhs)
    R=A_R-mean_rhs;
    L=B_L;
elseif abs(mean_lhs) > abs(mean_rhs)
    L=B_L-mean_lhs;
    R=A_R;
end
data_unfiltered=[R L];
% Regenerate the data by adding noise
[m,n]=size(data_unfiltered);
noise=randn(m,n)*1.5;    % Set the noise factor to 1.5
data_unfiltered_noise=data_unfiltered + noise;

% cut the data :rhs&lhs ***** 7680 = 30sec start time (time for
sample to become really relax)
%RL_cut=[R(7680:end,1) L(7680:end,1)];
%data_unfiltered=RL_cut;
%--------------------------------------------------------------------
-------------------------------
% manual artifact removal : data > 100micro & data <-100micro then
delete data for RHS&LHS combine
[i j]=find(data_unfiltered_noise>100);
data_unfiltered_noise(i,:)=[];
[i j]=find(data_unfiltered_noise<-100);
data_unfiltered_noise(i,:)=[];
%--------------------------------------------------------------------
-------------------------------
% Calculate length of signals
N=length(data_unfiltered_noise);
% Define the sampling frequency
Fs=256;
% Define the max frequency
Fs_Max=Fs/2;
% Define the frequency vector for plotting
freq=((1:N)*Fs)/(2*N);
ff=freq';
% Define the frequency band of raw data
dpass=[0.5 4];                  % Delta band
drange=range(dpass);            % Range of Delta band
tpass=[4 8];                    % Theta band
trange=range(tpass);            % Range of Theta band
apass=[8 13];                   % Alpha band
arange=range(apass);            % Range of Alpha band
bpass=[13 30];                  % Beta band
brange=range(bpass);            % Range of Beta band
```

```
totalpass=[0.5 30];             % Overall band
band=[dpass; tpass; apass; bpass];
numband=size(band,1);           % Number of band
cutoff=band/Fs_Max;
filter_order=74;
filter_type='bandpass';
N_Order=filter_order+1;
w1=window(@hamming,N_Order);

for j=1:numband
        %Calculate filter coefficient
        filter_coef=fir1(filter_order,cutoff(j,:),filter_type,w1);
        figure(j);
        for i=1:2           % for each channel (left & right)
        %Convolve the raw signal with filter coefficient

filt(j).out(:,i)=conv(data_unfiltered_noise(:,i),filter_coef);

        %Calculate FFT calculation
        fftsignal(j).out(:,i)=abs(fft(filt(j).out(:,i)));
        %Calculate the PSD from FFT
        psdsignal(j).out(:,i)=fftsignal(j).out(:,i).^2;
        p=length(fftsignal(j).out(:,i));
        %Calculate the signal frequency
        freq=(0:(p/2-1))*128/(p/2);
        %Calculate the mean frequency
        mpf1(i)=trapz(freq,freq.*psdsignal(j).out(1:p/2,i)') /
trapz(freq,psdsignal(j).out(1:p/2,i)');
        %Calculate Mean Power from PSD (Energy Spectral Density)
        area1(j,i)=trapz(freq,psdsignal(j).out(1:p/2,i)');
        freqrange1=band(j,2)- band(j,1);
        mp1=area1(j,i)/freqrange1;
        mpf1f(j,i)=mpf1(i);
        mp1f(j,i)=mp1;

        %Calculate the PSD using PSD function & window
        NFFT=1024;WINDOW=256;NOVERLAP=round(WINDOW/2); DFLAG='MEAN';
        [pxx,f]=psd(filt(j).out(:,i),NFFT,Fs,WINDOW,NOVERLAP,DFLAG);
        psdd(j).out(:,i)=pxx;
        %Calculate the mean frequency
        mpf2(i)=trapz(f,f.*pxx)/trapz(f,pxx);
        %Calculate the mean power from the PSD (ESD)
        area2(j,i)=trapz(f,pxx);
        freqrange2=band(j,2)- band(j,1);
        mp2=area2(j,i)/freqrange2;
        mpf2f(j,i)=mpf2(i);
        mp2f(j,i)=mp2;

        % Using Spectrum Welch
        h=spectrum.welch('Hamming',WINDOW,100*NOVERLAP/WINDOW);
        hpsd=psd(h,filt(j).out(:,i),'NFFT',NFFT,'Fs',Fs);
        hpsdd(j).out(:,i)=hpsd;
```

2. MATLAB Coding to Perform Spectral Centroids

```
function [C, CMean, CSD, CMax] =
SpecCentroid(snd,fs,nfft,window,noverlap)

%CENTROID Calculates spectral centroid of a sound.
%   C = CENRTROID(SOUND, FS) calculates the spectral centroid of a
given
%   sound. Fs is the sample frequency; the window length is the
minimum of 2048 and
%   the sample length; overlap is 80%.
%
%   [C, CMean, CSD, CMax] = CENRTROID(SOUND, FS) also calculates the
mean,
%   standard deviation and maximum of spectral centroid.
%
%   C = CENRTROID(SOUND,FS,NFFT);
%   C = CENRTROID(SOUND,FS,NFFT,WINDOW);
%   C = CENRTROID(SOUND,FS,NFFT,WINDOW,NOVERLAP)
%   Read the manual of the SPECGRAM command for the parameters.
%
% Example:
%   [s, fs] = wavread('testsound.wav');
%   c = centroid(s, fs);
% Uses:
%   Matlab Signal Processing Toolbox
%
% References:
%   Tzanetakis, G., Essl, G. and Cook, P.
%   In: Proc. Int. Symposium on Music Information
%   Retrieval (ISMIR), Bloomington, Indiana, 2001
%
% Frederik Nagel and Michael Großbach
% Institute of Music Physiology and Musicians' Medicine
% Hanover University of Music and Drama
% Hannover
% Germany
%
% e-mail: frederik.nagel@hmt-hannover.de
% homepage: http://www.immm.hmt-hannover.de
%
% May 29, 2006.
%
% See also SPECGRAM, FFT

error(nargchk(2, 5, nargin))
if(nargin==2)
    nfft = min([length(snd) 2048]);
    window = nfft;
    noverlap = round(window*.8);
    s = specgram(snd, nfft, fs, window, noverlap);
elseif (nargin==3)
    s = specgram(snd, nfft, fs);
elseif (nargin==4)
    s = specgram(snd, nfft, fs, window);
elseif (nargin==5)
    s = specgram(snd, nfft, fs, window, noverlap);
end
```

```
C = sum((repmat((1:size(s,1))',1,size(s,2)) .* abs(s))) ./
sum(abs(s));
CMean = mean(C);
CSD = std(C);
CMax = max(C);
```

3. **MATLAB Coding for the Classification of the selected EEG Asymmetry Features at 50:50 ratio Training and Testing**

```
%EEG DATA CLASSIFICATION USING K-NN ASYMMETRY FEATURES AT 50:50
TRAINING
%AND TESTING RATIO
clear all;
% Call the file containing the power ratio
load('Asymmetry.mat');
[m n]=size(data);

%Define the length of training group of data
G1_train=ones(25,1)*1.47; % EEG PSD data with questionnaires (EC
state)
G2_train=ones(25,1)*1.72; % EEG PSD data during IQ Test
G3_train=ones(20,1)*1.62; % EEG PSD data before HR
G4_train=ones(20,1)*1.46; % EEG PSD data after HR
Group_Class_Train=[G1_train;G2_train;G3_train;G4_train];
%Define the length of testing group of data
G1_test=ones(25,1)*1.47; % EEG PSD data with questionnaires (EC
state)
G2_test=ones(25,1)*1.72; % EEG PSD data during IQ Test
G3_test=ones(20,1)*1.62; % EEG PSD data before HR
G4_test=ones(20,1)*1.46; % EEG PSD data after HR
Group_Class_Test=[G1_test;G2_test;G3_test;G4_test];
%Define the Training & Testing Group at 50:50

%Training Group, 50%
Group_1_TR=data(1:25,2:5);
Group_2_TR=data(51:75,2:5);
Group_3_TR=data(101:120,2:5);
Group_4_TR=data(141:160,2:5);
Training_Data=[Group_1_TR; Group_2_TR; Group_3_TR; Group_4_TR];

%Testing Group, 50%
Group_1_Test=data(26:50,2:5);
Group_2_Test=data(76:100,2:5);
Group_3_Test=data(121:140,2:5);
Group_4_Test=data(161:180,2:5);
Testing_Data=[Group_1_Test; Group_2_Test; Group_3_Test;
Group_4_Test];

% Apply the k-NN classifier
for k=1:1:90
KNN_Output(:,k)=knnclassify(Testing_Data,Training_Data,Group_Class_Tr
ain,k,'euclidean','nearest');

[m1 n1]=size(Testing_Data);
for jj=1:m1
    if KNN_Output(jj,k)==Group_Class_Test(jj,1)
       KNN_result_compare(jj,k)=1;
    elseif KNN_Output(jj,k)~=Group_Class_Test(jj,1)
    KNN_result_compare(jj,k)=0;
    end
end
end
% Group percentage
% Group 1
KNN_G1=sum(KNN_result_compare(1:15,1));
```

```
KNN_G1_percentage=KNN_G1/15*100;
KNN_G1_PER=KNN_G1_percentage';
% Group 2
KNN_G2=sum(KNN_result_compare(16:30,1));
KNN_G2_percentage=KNN_G2/15*100;
KNN_G2_PER=KNN_G2_percentage';
% Group 3
KNN_G3=sum(KNN_result_compare(31:42,1));
KNN_G3_percentage=KNN_G3/12*100;
KNN_G3_PER=KNN_G3_percentage';
% Group 4
KNN_G4=sum(KNN_result_compare(43:54,1));
KNN_G4_percentage=KNN_G4/12*100;
KNN_G4_PER=KNN_G4_percentage';

% Overall group percentage
KNN_result_compare_sum=sum(KNN_result_compare);
KNN_result_compare_percentage=KNN_result_compare_sum/m1*100;
KNN_PER=KNN_result_compare_percentage';
```

4. MATLAB Coding to Classify the selected EEG Asymmetry Features at 70:30 ratio Training and Testing

```
%EEG DATA CLASSIFICATION USING K-NN ASYMMETRY FEATURES AT 70:30
TRAINING AND %TESTING RATIO
clear all;
% Call the file containing the normalization data
load('Asymmetry.mat');
[m n]=size(data);

%Define the length of training group of data
G1_train=ones(35,1)*1.47; % EEG PSD data with questionnaires (EC
state)
G2_train=ones(35,1)*1.72; % EEG PSD data during IQ Test
G3_train=ones(28,1)*1.62; % EEG PSD data before HR
G4_train=ones(28,1)*1.46; % EEG PSD data after HR
Group_Class_Train=[G1_train;G2_train;G3_train;G4_train];
%Define the length of testing group of data
G1_test=ones(15,1)*1.47; % EEG PSD data with questionnaires (EC
state)
G2_test=ones(15,1)*1.72; % EEG PSD data during IQ Test
G3_test=ones(12,1)*1.62; % EEG PSD data before HR
G4_test=ones(12,1)*1.46; % EEG PSD data after HR
Group_Class_Test=[G1_test;G2_test;G3_test;G4_test];
%Define the Training & Testing Group at 70:30

%Training Group, 70%
Group_1_TR=data(1:35,2:5);
Group_2_TR=data(51:85,2:5);
Group_3_TR=data(101:128,2:5);
Group_4_TR=data(141:168,2:5);
Training_Data=[Group_1_TR; Group_2_TR; Group_3_TR; Group_4_TR];

%Testing Group, 30%
Group_1_Test=data(36:50,2:5);
Group_2_Test=data(86:100,2:5);
Group_3_Test=data(129:140,2:5);
Group_4_Test=data(169:180,2:5);
Testing_Data=[Group_1_Test; Group_2_Test; Group_3_Test;
Group_4_Test];

% Apply the k-NN classifier
for k=1:1:126
KNN_Output(:,k)=knnclassify(Testing_Data,Training_Data,Group_Class_Tr
ain,k,'euclidean','nearest');
[m1 n1]=size(Testing_Data);
for jj=1:m1
    if KNN_Output(jj,k)==Group_Class_Test(jj,1)
       KNN_result_compare(jj,k)=1;
    elseif KNN_Output(jj,k)~=Group_Class_Test(jj,1)
    KNN_result_compare(jj,k)=0;
    end
end
end
% Group percentage
% Group 1
KNN_G1=sum(KNN_result_compare(1:15,1));
KNN_G1_percentage=KNN_G1/15*100;
KNN_G1_PER=KNN_G1_percentage';
% Group 2
```

```
KNN_G2=sum(KNN_result_compare(16:30,1));
KNN_G2_percentage=KNN_G2/15*100;
KNN_G2_PER=KNN_G2_percentage';
% Group 3
KNN_G3=sum(KNN_result_compare(31:42,1));
KNN_G3_percentage=KNN_G3/12*100;
KNN_G3_PER=KNN_G3_percentage';
% Group 4
KNN_G4=sum(KNN_result_compare(43:54,1));
KNN_G4_percentage=KNN_G4/12*100;
KNN_G4_PER=KNN_G4_percentage';

% Overall group percentage
KNN_result_compare_sum=sum(KNN_result_compare);
KNN_result_compare_percentage=KNN_result_compare_sum/m1*100;
KNN_PER=KNN_result_compare_percentage';
```

5. **Classification of the selected EEG Natural Log Asymmetry Features at 50:50 ratio Training and Testing**

```
%EEG DATA CLASSIFICATION USING K-NN NATURAL LOG ASYMMETRY AT 50:50
TRAINING
%AND TESTING RATIO
clear all;
% Call the file containing the power ratio
load('knn_log_asymmetry.mat');
[m n]=size(data);

%Define the length of training group of data
G1_train=ones(25,1)*1.3;  % EEG PSD data with questionnaires (EC
state)
G2_train=ones(25,1)*1.25; % EEG PSD data during IQ Test
G3_train=ones(20,1)*1.25; % EEG PSD data before HR
G4_train=ones(20,1)*1.3;  % EEG PSD data after HR
Group_Class_Train=[G1_train;G2_train;G3_train;G4_train];
%Define the length of testing group of data
G1_test=ones(25,1)*1.3;  % EEG PSD data with questionnaires (EC
state)
G2_test=ones(25,1)*1.25; % EEG PSD data during IQ Test
G3_test=ones(20,1)*1.25; % EEG PSD data before HR
G4_test=ones(20,1)*1.3;  % EEG PSD data after HR
Group_Class_Test=[G1_test;G2_test;G3_test;G4_test];
%Define the Training & Testing Group at 50:50

%Training Group, 50%
Group_1_TR=data(1:25,2:5);
Group_2_TR=data(51:75,2:5);
Group_3_TR=data(101:120,2:5);
Group_4_TR=data(141:160,2:5);
Training_Data=[Group_1_TR; Group_2_TR; Group_3_TR; Group_4_TR];

%Testing Group, 50%
Group_1_Test=data(26:50,2:5);
Group_2_Test=data(76:100,2:5);
Group_3_Test=data(121:140,2:5);
Group_4_Test=data(161:180,2:5);
Testing_Data=[Group_1_Test; Group_2_Test; Group_3_Test;
Group_4_Test];

% Apply the k-NN classifier
for k=1:1:90
KNN_Output(:,k)=knnclassify(Testing_Data,Training_Data,Group_Class_Tr
ain,k,'euclidean','nearest');

[m1 n1]=size(Testing_Data);
for jj=1:m1
    if KNN_Output(jj,k)==Group_Class_Test(jj,1)
       KNN_result_compare(jj,k)=1;
    elseif KNN_Output(jj,k)~=Group_Class_Test(jj,1)
    KNN_result_compare(jj,k)=0;
    end
end
end
% Group percentage
% Group 1
KNN_G1=sum(KNN_result_compare(1:15,1));
```

```
KNN_G1_percentage=KNN_G1/15*100;
KNN_G1_PER=KNN_G1_percentage';
% Group 2
KNN_G2=sum(KNN_result_compare(16:30,1));
KNN_G2_percentage=KNN_G2/15*100;
KNN_G2_PER=KNN_G2_percentage';
% Group 3
KNN_G3=sum(KNN_result_compare(31:42,1));
KNN_G3_percentage=KNN_G3/12*100;
KNN_G3_PER=KNN_G3_percentage';
% Group 4
KNN_G4=sum(KNN_result_compare(43:54,1));
KNN_G4_percentage=KNN_G4/12*100;
KNN_G4_PER=KNN_G4_percentage';

% Overall group percentage
KNN_result_compare_sum=sum(KNN_result_compare);
KNN_result_compare_percentage=KNN_result_compare_sum/m1*100;
KNN_PER=KNN_result_compare_percentage';
```

6. **Classification of the selected EEG Natural Log Asymmetry Features at 70:30 ratio Training and Testing**

```
%EEG DATA CLASSIFICATION USING K-NN NATURAL LOG ASYMMETRY FEATURE AT
70:30
%TRAINING AND TESTING RATIO
clear all;
% Call the file containing the normalization data
load('knn_log_asymmetry.mat');
[m n]=size(data);

%Define the length of training group of data
G1_train=ones(35,1)*1.3;  % EEG PSD data with questionnaires (EC
state)
G2_train=ones(35,1)*1.25; % EEG PSD data during IQ Test
G3_train=ones(28,1)*1.25; % EEG PSD data before HR
G4_train=ones(28,1)*1.3;  % EEG PSD data after HR
Group_Class_Train=[G1_train;G2_train;G3_train;G4_train];
%Define the length of testing group of data
G1_test=ones(15,1)*1.3;  % EEG PSD data with questionnaires (EC
state)
G2_test=ones(15,1)*1.25; % EEG PSD data during IQ Test
G3_test=ones(12,1)*1.25; % EEG PSD data before HR
G4_test=ones(12,1)*1.3;  % EEG PSD data after HR
Group_Class_Test=[G1_test;G2_test;G3_test;G4_test];
%Define the Training & Testing Group at 70:30

%Training Group, 70%
Group_1_TR=data(1:35,2:5);
Group_2_TR=data(51:85,2:5);
Group_3_TR=data(101:128,2:5);
Group_4_TR=data(141:168,2:5);
Training_Data=[Group_1_TR; Group_2_TR; Group_3_TR; Group_4_TR];

%Testing Group, 30%
Group_1_Test=data(36:50,2:5);
Group_2_Test=data(86:100,2:5);
Group_3_Test=data(129:140,2:5);
Group_4_Test=data(169:180,2:5);
Testing_Data=[Group_1_Test; Group_2_Test; Group_3_Test;
Group_4_Test];

% Apply the k-NN classifier
for k=1:1:126
KNN_Output(:,k)=knnclassify(Testing_Data,Training_Data,Group_Class_Tr
ain,k,'euclidean','nearest');

[m1 n1]=size(Testing_Data);
for jj=1:m1
    if KNN_Output(jj,k)==Group_Class_Test(jj,1)
       KNN_result_compare(jj,k)=1;
    elseif KNN_Output(jj,k)~=Group_Class_Test(jj,1)
    KNN_result_compare(jj,k)=0;
    end
end
end
% Group percentage
% Group 1
KNN_G1=sum(KNN_result_compare(1:15,1));
```

```
KNN_G1_percentage=KNN_G1/15*100;
KNN_G1_PER=KNN_G1_percentage';
% Group 2
KNN_G2=sum(KNN_result_compare(16:30,1));
KNN_G2_percentage=KNN_G2/15*100;
KNN_G2_PER=KNN_G2_percentage';
% Group 3
KNN_G3=sum(KNN_result_compare(31:42,1));
KNN_G3_percentage=KNN_G3/12*100;
KNN_G3_PER=KNN_G3_percentage';
% Group 4
KNN_G4=sum(KNN_result_compare(43:54,1));
KNN_G4_percentage=KNN_G4/12*100;
KNN_G4_PER=KNN_G4_percentage';

% Overall group percentage
KNN_result_compare_sum=sum(KNN_result_compare);
KNN_result_compare_percentage=KNN_result_compare_sum/m1*100;
KNN_PER=KNN_result_compare_percentage';
```

7. Classification of the selected EEG RER Features at 50:50 ratio Training and Testing

```
%EEG DATA CLASSIFICATION USING K-NN USING RELATIVE ENERGY RATIO (RER)
WITH
%50:50 RATIO
clear all;
% Call the file containing the normalization data
load('knn_power_ratio.mat');
[m n]=size(data);

%Define the length of training group of data
G1_train=ones(25,1)*1.53; % EEG PSD data with questionnaires (EC
state)
G2_train=ones(25,1)*1.23; % EEG PSD data during IQ Test
G3_train=ones(20,1)*1.49; % EEG PSD data before HR
G4_train=ones(20,1)*1.46; % EEG PSD data after HR
Group_Class_Train=[G1_train;G2_train;G3_train;G4_train];

%Define the length of testing group of data
G1_test=ones(25,1)*1.53; % EEG PSD data with questionnaires (EC
state)
G2_test=ones(25,1)*1.23; % EEG PSD data during IQ Test
G3_test=ones(20,1)*1.49; % EEG PSD data before HR
G4_test=ones(20,1)*1.46; % EEG PSD data after HR
Group_Class_Test=[G1_test;G2_test;G3_test;G4_test];

%Training Group, 50%
Group_1_TR=data(1:25,2:5);
Group_2_TR=data(51:75,2:5);
Group_3_TR=data(101:120,2:5);
Group_4_TR=data(141:160,2:5);
Training_Data=[Group_1_TR; Group_2_TR; Group_3_TR; Group_4_TR];

%Testing Group, 50%
Group_1_Test=data(26:50,2:5);
Group_2_Test=data(76:100,2:5);
Group_3_Test=data(121:140,2:5);
Group_4_Test=data(161:180,2:5);
Testing_Data=[Group_1_Test; Group_2_Test; Group_3_Test;
Group_4_Test];

% Apply the k-NN classifier
for k=1:1:90
KNN_Output(:,k)=knnclassify(Testing_Data,Training_Data,Group_Class_Tr
ain,k,'euclidean','nearest');

[m1 n1]=size(Testing_Data);
for jj=1:m1
    if KNN_Output(jj,k)==Group_Class_Test(jj,1)
       KNN_result_compare(jj,k)=1;
    elseif KNN_Output(jj,k)~=Group_Class_Test(jj,1)
    KNN_result_compare(jj,k)=0;
    end
end
end
% Group percentage
% Group 1
KNN_G1=sum(KNN_result_compare(1:15,2:5));
KNN_G1_percentage=KNN_G1/15*100;
```

```
KNN_G1_PER=KNN_G1_percentage';
% Group 2
KNN_G2=sum(KNN_result_compare(16:30,2:5));
KNN_G2_percentage=KNN_G2/15*100;
KNN_G2_PER=KNN_G2_percentage';
% Group 3
KNN_G3=sum(KNN_result_compare(31:42,2:5));
KNN_G3_percentage=KNN_G3/12*100;
KNN_G3_PER=KNN_G3_percentage';
% Group 4
KNN_G4=sum(KNN_result_compare(43:54,2:5));
KNN_G4_percentage=KNN_G4/12*100;
KNN_G4_PER=KNN_G4_percentage';

% Overall group percentage
KNN_result_compare_sum=sum(KNN_result_compare);
KNN_result_compare_percentage=KNN_result_compare_sum/m1*100;
KNN_PER=KNN_result_compare_percentage';
```

8. Classification of the selected EEG RER Features at 70:30 ratio Training and Testing

```
%EEG DATA CLASSIFICATION USING K-NN USING RELATIVE ENERGY RATIO (RER)
at
%70:30 training and testing ratio
clear all;
% Call the file containing the normalization data
load('knn_power_ratio.mat');
[m n]=size(data);

%Define the length of training group of data with the centroids of
RER
G1_train=ones(35,1)*1.53; % EEG PSD data with questionnaires at EC
state
G2_train=ones(35,1)*1.23; % EEG PSD data during IQ Test
G3_train=ones(28,1)*1.49; % EEG PSD data before HR
G4_train=ones(28,1)*1.46; % EEG PSD data after HR
Group_Class_Train=[G1_train;G2_train;G3_train;G4_train];

%Define the length of testing group of data with the centroids of RER
G1_test=ones(15,1)*1.53; % EEG PSD data with questionnaires at EC
state
G2_test=ones(15,1)*1.23; % EEG PSD during IQ Test
G3_test=ones(12,1)*1.49; % EEG PSD before HR
G4_test=ones(12,1)*1.46; % EEG PSD after HR
Group_Class_Test=[G1_test;G2_test;G3_test;G4_test];

%Define the Training & Testing Group at 70:30

%Training Group, 70%
Group_1_TR=data(1:35,2:5);
Group_2_TR=data(51:85,2:5);
Group_3_TR=data(101:128,2:5);
Group_4_TR=data(141:168,2:5);
Training_Data=[Group_1_TR; Group_2_TR; Group_3_TR; Group_4_TR];

%Testing Group, 30%
Group_1_Test=data(36:50,2:5);
Group_2_Test=data(86:100,2:5);
Group_3_Test=data(129:140,2:5);
Group_4_Test=data(169:180,2:5);
Testing_Data=[Group_1_Test; Group_2_Test; Group_3_Test;
Group_4_Test];

% Apply the k-NN classifier
for k=1:1:126
KNN_Output(:,k)=knnclassify(Testing_Data,Training_Data,Group_Class_Tr
ain,k,'euclidean','nearest');

[m1 n1]=size(Testing_Data);
for jj=1:m1
    if KNN_Output(jj,k)==Group_Class_Test(jj,1)
       KNN_result_compare(jj,k)=1;
    elseif KNN_Output(jj,k)~=Group_Class_Test(jj,1)
    KNN_result_compare(jj,k)=0;
    end
end
end
```

```
% Group percentage
% Group 1
KNN_G1=sum(KNN_result_compare(1:15,2:5));
KNN_G1_percentage=KNN_G1/15*100;
KNN_G1_PER=KNN_G1_percentage';
% Group 2
KNN_G2=sum(KNN_result_compare(16:30,2:5));
KNN_G2_percentage=KNN_G2/15*100;
KNN_G2_PER=KNN_G2_percentage';
% Group 3
KNN_G3=sum(KNN_result_compare(31:42,2:5));
KNN_G3_percentage=KNN_G3/12*100;
KNN_G3_PER=KNN_G3_percentage';
% Group 4
KNN_G4=sum(KNN_result_compare(43:54,2:5));
KNN_G4_percentage=KNN_G4/12*100;
KNN_G4_PER=KNN_G4_percentage';

% Overall group percentage
KNN_result_compare_sum=sum(KNN_result_compare);
KNN_result_compare_percentage=KNN_result_compare_sum/m1*100;
KNN_PER=KNN_result_compare_percentage';
```

9. Confusion Matrix

```
function [confmatrix] = cfmatrix(actual, predict, classlist, per)
% CFMATRIX calculates the confusion matrix for any prediction
% algorithm that generates a list of classes to which the test
% feature vectors are assigned
%
% Outputs: confusion matrix
%
%                  Actual Classes
%                    p       n
%               ___|_____|______|
%    Predicted  p'|     |      |
%      Classes  n'|     |      |
%
% Inputs:
% 1. actual / 2. predict
% The inputs provided are the 'actual' classes vector
% and the 'predict'ed classes vector. The actual classes are the
classes
% to which the input feature vectors belong. The predicted classes
are the
% class to which the input feature vectors are predicted to belong
to,
% based on a prediction algorithm.
% The length of actual class vector and the predicted class vector
need to
% be the same. If they are not the same, an error message is
displayed.
% 3. classlist
% The third input provides the list of all the classes {p,n,...} for
which
% the classification is being done. All classes are numbers.
% 4. per = 1/0 (default = 0)
% This parameter when set to 1 provides the values in the confusion
matrix
% as percentages. The default provides the values in numbers.
%
% Example:
% >> a = [ 1 2 3 1 2 3 1 1 2 3 2 1 1 2 3];
% >> b = [ 1 2 3 1 2 3 1 1 1 2 2 1 2 1 3];
% >> Cf = cfmatrix(a, b);
%
% [Avinash Uppuluri: avinash_uv@yahoo.com: Last modified: 08/21/08]

% If classlist not entered: make classlist equal to all
% unique elements of actual
if (nargin < 2)
   error('Not enough input arguments.');
elseif (nargin == 2)
    classlist = unique(actual); % default values from actual
    per = 0; % default is numbers and input 1 for percentage
elseif (nargin == 3)
    per = 0; % default is numbers and input 1 for percentage
end

if (length(actual) ~= length(predict))
    error('First two inputs need to be vectors with equal size.');
elseif ((size(actual,1) ~= 1) && (size(actual,2) ~= 1))
```

```
    error('First input needs to be a vector and not a matrix');
elseif ((size(predict,1) ~= 1) && (size(predict,2) ~= 1))
    error('Second input needs to be a vector and not a matrix');
end
format short g;
n_class = length(classlist);
line_two = '----------';
line_three = '_________|';
for i = 1:n_class
    obind_class_i = find(actual == classlist(i));
    prind_class_i = find(predict == classlist(i));
    confmatrix(i,i) = length(intersect(obind_class_i,prind_class_i));
    for j = 1:n_class
        %if (j ~= i)
        if (j < i)
        % observed j predicted i
        confmatrix(i,j) = length(find(actual(prind_class_i) ==
classlist(j)));
        % observed i predicted j
        confmatrix(j,i) = length(find(predict(obind_class_i) ==
classlist(j)));
        end
    end
    line_two = strcat(line_two,'---',num2str(classlist(i)),'-----');
    line_three = strcat(line_three,'__________');
end

if (per == 1)
    confmatrix = (confmatrix ./ length(actual)).*100;
end
% output to screen
disp('------------------------------------------');
disp('             Actual Classes');
disp(line_two);
disp('Predicted|                     ');
disp('  Classes|                     ');
disp(line_three);
for i = 1:n_class
    temps = sprintf('      %d       ',i);
    for j = 1:n_class
    temps = strcat(temps,sprintf(' |   %2.1f    ',confmatrix(i,j)));
    end
    disp(temps);
    clear temps
end
disp('------------------------------------------');
%TP,FP,FN,TN
% True Postive [TP] = Condition Present + Positive result
% False Positive [FP] = Condition absent + Positive result [Type
% I error]
% False (invalid) Negative [FN] = Condition present + Negative result
[Type
% II error]
% True (accurate) Negative [TN] = Condition absent + Negative result
disp('--------------------------------------');
disp(' Actual Classes');
disp(line_two);
temps = sprintf(' TP ');
for i = 1:n_class
    temps = strcat(temps,sprintf(' | %2.1f ',confmatrix(i,i)));
end
```

```
disp(temps);
clear temps
temps = sprintf(' FP ');
for i = 1:n_class
    temps = strcat(temps,sprintf(' | %2.1f ',sum(confmatrix(i,:))-
confmatrix(i,i) ));
end
disp(temps);
clear temps
temps = sprintf(' FN ');
for i = 1:n_class
    temps = strcat(temps,sprintf(' | %2.1f ',sum(confmatrix(:,i))-
confmatrix(i,i) ));
end
disp(temps);
clear temps
temps = sprintf(' TN ');
for i = 1:n_class
    temps = strcat(temps,sprintf(' | %2.1f ',sum(diag(confmatrix))-
confmatrix(i,i) ));
end
disp(temps);
clear temps
temps =sprintf('SEN');
for i=1:n_class
    temps = strcat(temps,sprintf(' |
%2.4f',(confmatrix(i,i)/(confmatrix(i,i)+sum(confmatrix(:,i))-
confmatrix(i,i)))));
end
disp(temps);
clear temps
temps=sprintf('SPEC');
for i=1:n_class
    temps = strcat(temps,sprintf(' | %2.4f',(sum(diag(confmatrix))-
confmatrix(i,i))/(sum(diag(confmatrix))-
confmatrix(i,i)+sum(confmatrix(i,:))-confmatrix(i,i))));
end
disp(temps);
clear temps
```

10. Cross-validation of the k-NN Classifier using FKM

```
% CROSS VALIDATION OF EEG FEATURES USING FKM AND CROSSVALIND
clear all;
%Load EEG features
load('knn_power_ratio.mat');
a2=data(1:180,2:5);
opts = statset('Display','final');
[idx,ctrs] =
kmeans(a2,4,'Distance','sqEuclidean','Replicates',5,'Options',opts);

for z=1:1:4
sum1(z,:)=sum(idx(1:50,1)==z);
sum2(z,:)=sum(idx(51:100,1)==z);
sum3(z,:)=sum(idx(101:140,1)==z);
sum4(z,:)=sum(idx(141:180,1)==z);
end
%sum1,sum2,sum3,sum4==group kita #1,#2,#3,#4
      %g1  g2  g3  g4 >> group kmeans
%sum1  11  10  23  6
%sum2  10  40   0  0
%sum3  12  11  12  5
%sum4  14  9   12  5
% select the group
% group kita #1 >> 23(g3)
% group kita #2 >> 40(g2)
% group kita #3 >> 12(g1)
% group kita #4 >> 5(g4)
%Re-assign the groups
%data group 1:50 >> kmeans group #3 so b1==3
%data group 51:100 >> kmeans group #2 so b2==2
%data group 101:140 >> kmeans group #1 so b1==1
%data group 141:180 >> kmeans group #4 so b2==4
b1=3;b2=2;b3=1;b4=4;
A=find(idx(1:50,:)==b1);[mm1,nn1]=size(A);
B=find(idx(51:100,:)==b2);[mm2,nn2]=size(B);
C=find(idx(101:140,:)==b3);[mm3,nn3]=size(C);
D=find(idx(141:180,:)==b4);[mm4,nn4]=size(D);
E=size(A)+size(B)+size(C)+size(D);
g1=5;g2=40;g3=20;g4=5;

%Relate the EEG data according to the FKM index to produce FKM new
groups
X_NEW=[A;B;C;D];
data_kmean=a2(X_NEW,:);
[k,l]=size(data_kmean);
group=[repmat(1,mm1,1);repmat(2,mm2,1);repmat(3,mm3,1);repmat(4,mm4,1
)];
%Define the number of fold according to the k value of the FKM
fold=ceil(k/10);
indices = crossvalind('Kfold',group,fold);

%cross-validation process k-fold >> to check training data validity
>> target 100% accuracy

for iii = 1:fold
test3(:,iii) = (indices == iii);
train3(:,iii) = ~test3 (:,iii);
```

```
Class(:,iii)=
knnclassify(data_kmean(test3(:,iii),:),data_kmean(train3(:,iii),:),gr
oup(train3(:,iii),:));
test_g(:,iii)=group(test3(:,iii)~=0,1);

end
[m n] = size(Class);
for j=1:m
if Class(j)==test_g(j)
    result(j)=1;
elseif Class(j)~=test_g(j)
    result(j)=0;
end
end
%check the performance of k-fold cross-validation
total_percentage=mean(result)*100;
per_av=mean(mean(total_percentage,2));
```

11. Cross-validation of the k-NN Classifier without using FKM

```
%EEG DATA CLASSIFICATION USING K-NN : k-Fold Cross Validation
clear all;
% Call the file containing the normalization data
load('knn_power_ratio.mat');
[m n]=size(data);

%Define the length of training group of data
G1_train=ones(45,1)*1.53; % EEG PSD data with questionnaires (EC
state)
G2_train=ones(45,1)*1.23; % EEG PSD data during IQ Test data
G3_train=ones(32,1)*1.49; % EEG PSD data before HR
G4_train=ones(32,1)*1.46; % EEG PSD data after HR
Group_Class_Train=[G1_train;G2_train;G3_train;G4_train];

%Define the length of testing group of data
G1_test=ones(5,1)*1.53; % EEG PSD data with questionnaires (EC state)
G2_test=ones(5,1)*1.23; % EEG PSD data during IQ Test
G3_test=ones(8,1)*1.49; % EEG PSD data before HR
G4_test=ones(8,1)*1.46; % EEG PSD data after HR
Group_Class_Test=[G1_test;G2_test;G3_test;G4_test];

%Define the Training & Testing data for 10-fold cross validation
test_data1=data(1:5,2:5);
test_data2=data(51:55,2:5);
test_data3=data(101:108,2:5);
test_data4=data(141:148,2:5);

Tr_1=data(6:50,2:5);
Tr_2=data(56:100,2:5);
Tr_3=data(109:140,2:5);
Tr_4=data(149:180,2:5);
Testing_data=[test_data1;test_data2;test_data3;test_data4];
Training_data=[Tr_1;Tr_2;Tr_3;Tr_4];

% Apply the k-NN classifier
for k=1:1:50
KNN_Output(:,k)=knnclassify(Testing_data,Training_data,Group_Class_Tr
ain,k,'euclidean','nearest');

[m1 n1]=size(Testing_data);
for jj=1:m1
    if KNN_Output(jj,k)==Group_Class_Test(jj,1)
       KNN_result_compare(jj,k)=1;
    elseif KNN_Output(jj,k)~=Group_Class_Test(jj,1)
    KNN_result_compare(jj,k)=0;
    end
end
end

% Overall group percentage
KNN_result_compare_sum=sum(KNN_result_compare);
KNN_result_compare_percentage=KNN_result_compare_sum/m1*100;
KNN_PER=KNN_result_compare_percentage';

data=KNN_Output;
[m2 n2]=size(data);
for b=1:1:n2
```

```
for a=1:1:m2

    if data(a,b)==1.23
        data2(a,b)=1;
    elseif data(a,b)==1.46
        data2(a,b)=2;
    elseif data(a,b)==1.49
        data2(a,b)=3;
    elseif data(a,b)==1.53
        data2(a,b)=4;
    end
end
end

data3=Group_Class_Test;
[m n]=size(data3);
b=1;
for a=1:1:m

    if data3(a,b)==1.23
        data22(a,b)=1;
    elseif data3(a,b)==1.46
        data22(a,b)=2;
    elseif data3(a,b)==1.49
        data22(a,b)=3;
    elseif data3(a,b)==1.53
        data22(a,b)=4;
    end
end

% Measure the performance of the cross validation
class_perf=classperf(data2(:,1),data22);
performance_CorrectRate=class_perf.CorrectRate;
performance_Sensitivity=class_perf.Sensitivity;
performance_Specificity=class_perf.Specificity;
```

12. Cross-validation of the k-NN Classifier using LOO

```
%CROSS-VALIDATION OF k-NN CLASSIFICATION USING LEAVE-ONE-OUT (LOO)
%LOL for k-NN
%g1,.....=nth group no of observations;
%data_g=[g1;g2;g3;......];
%name your data as "a"
%define the group
g1=11;g2=10;g3=23;g4=6;
data_g=[g1;g2;g3;g4];
data_g_sort=sort(data_g);
fold=data_g_sort(1,1);

group=[repmat(1,g1,1); repmat(2,g2,1); repmat(3,g3,1);
repmat(4,g4,1)];

for i=1:1:fold
    a=data(1:180,2:5);
    data_train_test=a;
    data_class=group;

    t1=i;
    t2=i+g1;
    t3=i+g1+g2;
    t4=i+g1+g2+g3;
    t_all=[t1;t2;t3;t4];

    test_data(i).out=data_train_test(t_all,:);
    class_test(i).out=data_class(t_all,:);

    data_train_test(t_all,:)=[];
    train_data(i).out=data_train_test;
    data_class(t_all,:)=[];
    class(i).out=data_class;

end
for TT=1:1:fold

    test_data2=test_data(TT).out;
    train_data2=train_data(TT).out;
    class2=class(TT).out;
    class_test2=class_test(TT).out;

    [m1 n1]=size(test_data2);
for k=1:1:10 % put your max K-values here

KNN(TT).out(:,k)=knnclassify(test_data2,train_data2,class2,k,'euclide
an','nearest');

for jj=1:m1
if KNN(TT).out(jj,k)==class_test2(jj,1)
        KNN_result_compare(TT).out(jj,k)=1;
elseif KNN(TT).out(jj,k)~=class_test2(jj,1)
        KNN_result_compare(TT).out(jj,k)=0;
end
    end
end
KNN_result_compare_sum(TT).out=sum(KNN_result_compare(TT).out);
```

```
KNN_result_compare_percentage(TT).out=(KNN_result_compare_sum(TT).out
/m1)*100;
 KNN_sum(TT,:)=KNN_result_compare_percentage(TT).out;
 end
KNN_av_LOL=mean(KNN_sum);
```

13. FCM Clustering

```
%FCM CLUSTERING OF EEG FEATURES
clear all;
% Call the file containing the power ratio data
load('knn_power_ratio.mat');
[m n]=size(data);
%Apply Fuzzy C-Mean with 4 number of cluster
datax=data(1:180,2:5);
[center,U,objFcn] = fcm(datax,4,NaN);
maxU = max(U);
index1 = find(U(1, :) == maxU);
index2 = find(U(2, :) == maxU);
index3 = find(U(3, :) == maxU);
index4 = find(U(4, :) == maxU);

%Plot the maximum membership grade, U
MAXU=maxU';
hist(MAXU(1:180,1));
% Create xlabel
xlabel('Grade of membership, U','FontName','Times New
Roman','fontsize',10,'fontweight','b');
%xlabel('Grade of Membership, U');

% Create ylabel
ylabel('Number of membership grade','FontName','Times New
Roman','fontsize',10,'fontweight','b');
%ylabel('Number of membership grade');

% Create title
%title('Histogram plot of FCM Membership Grade');
title('Histogram plot of FCM Membership Grade','FontName','Times New
Roman','fontsize',10,'fontweight','b');

% Calculate Centroids for each cluster using index obtained from
% membership, U
ind1=index1';[m1,n1]=size(ind1);
ind2=index2';[m2,n2]=size(ind2);
ind3=index3';[m3,n3]=size(ind3);
ind4=index4';[m4,n4]=size(ind4);
dataind1=datax(ind1,1); [mm1,nn1]=size(dataind1);
dataind2=datax(ind2,2); [mm2,nn2]=size(dataind2);
dataind3=datax(ind3,3); [mm3,nn3]=size(dataind3);
dataind4=datax(ind4,4); [mm4,nn4]=size(dataind4);
% apply centroid to indexed data with index size
cenind1=centroid(dataind1,m1);
cenind2=centroid(dataind2,m2);
cenind3=centroid(dataind3,m3);
cenind4=centroid(dataind4,m4);

% Plot the Iteration graph
figure
plot(objFcn)
title('Objective Function Values')
xlabel('Iteration Count')
ylabel('Objective Function Value')

% Plot the FCM for all clusters
figure
```

```
line(datax(ind1, 1), datax(ind1, 2),'linestyle',...
'none','marker', '+','color','b');

line(datax(ind2,1),datax(ind2,2),'linestyle',...
'none','marker', '*','color','y');

line(datax(ind3, 3), datax(ind3, 4),'linestyle',...
'none','marker', 'o','color','g');

line(datax(ind4,3),datax(ind4,4),'linestyle',...
'none','marker', 'x','color','r');

title('FCM Clustering','FontName','Times New
Roman','fontsize',10,'fontweight','b');
hold on
plot(center(1,1),center(1,2),'ko','markersize',15,'LineWidth',2)
plot(center(2,1),center(2,2),'ko','markersize',15,'LineWidth',2)
plot(center(3,3),center(3,4),'ko','markersize',15,'LineWidth',2)
plot(center(4,3),center(4,4),'ko','markersize',15,'LineWidth',2)
legend('Cluster 1','Cluster 2','Cluster 3','Cluster
4','Centroids','Location','NW')
legend ('show');
set(legend,'Orientation','horizontal','Location','SouthOutside',...
    'FontName','Times New Roman','fontsize',10,'fontweight','b');
%set(legend,'Location','SouthOutside');
hold off
Center1_val=maxU(1,index1);
Center2_val=maxU(1,index2);
Center3_val=maxU(1,index3);
Center4_val=maxU(1,index4);
```

14. FKM Clustering

```
%CLUSTERING THE EEG FEATURES USING FKM
clear all;
GrA=50;%size group A
GrB=50;%size group B
GrC=40;%size group C
GrD=40;%size group D

% Call the file containing the power ratio data
load('knn_power_ratio.mat');
[m n]=size(data);
%Apply Fuzzy C-Mean with 4 number of cluster
datax=data(1:180,2:5);
title('FKM Clustering','FontName','Times New
Roman','fontsize',10,'fontweight','b');
%hold on
opts = statset('Display','final');
[idx,ctrs] =
kmeans(datax,4,'Distance','sqEuclidean','Replicates',5,'Options',opts
);
subplot(1,2,1)
plot(datax(idx==1,1),datax(idx==1,2),'y.','MarkerSize',12);
hold on
plot(datax(idx==2,1),datax(idx==2,2),'r.','MarkerSize',12)
hold on
plot(datax(idx==3,3),datax(idx==3,4),'b.','MarkerSize',12)
hold on
plot(datax(idx==4,3),datax(idx==4,4),'g.','MarkerSize',12)

plot(ctrs(1,1),ctrs(1,2),'kx','MarkerSize',12,'LineWidth',2)
plot(ctrs(2,1),ctrs(2,2),'kx','MarkerSize',12,'LineWidth',2)
plot(ctrs(3,3),ctrs(3,4),'kx','MarkerSize',12,'LineWidth',2)
plot(ctrs(4,3),ctrs(4,4),'kx','MarkerSize',12,'LineWidth',2)
title('FKM Clustering','FontName','Times New
Roman','fontsize',10,'fontweight','b');
legend('Cluster 1','Cluster 2','Cluster 3','Cluster
4','Centroids','Location','NW')
legend ('show');
set(legend,'Orientation','vertical','Location','SouthOutside',...
    'FontName','Times New Roman','fontsize',10,'fontweight','b');
%set(legend,'Location','SouthOutside');
hold off

A_idx=find(idx(1:GrA,:)==1);
B_idx=find(idx(GrA+1:GrA+GrB,:)==2);
C_idx=find(idx(GrA+GrB+1:GrA+GrB+GrC,:)==3);
D_idx=find(idx(GrA+GrB+GrC+1:GrA+GrB+GrC+GrD,:)==4);
%
A=datax(A_idx,:);[mm1,nn1]=size(A);
B=datax(B_idx,:);[mm2,nn2]=size(B);
C=datax(C_idx,:);[mm3,nn3]=size(C);
D=datax(D_idx,:);[mm4,nn4]=size(D);
%
X_NEW=[A;B;C;D];
x2=X_NEW;
opts2 = statset('Display','final');
[idx2,ctrs2] =
kmeans(x2,4,'Distance','sqEuclidean','Replicates',5,'Options',opts);
```

```
subplot(1,2,2)
plot(x2(idx2==1,1),x2(idx2==1,2),'y.','MarkerSize',12)
hold on
plot(x2(idx2==2,1),x2(idx2==2,2),'r.','MarkerSize',12)
hold on
plot(x2(idx2==3,3),x2(idx2==3,4),'b.','MarkerSize',12)
hold on
plot(x2(idx2==4,3),x2(idx2==4,4),'g.','MarkerSize',12)
%
plot(ctrs2(1,1),ctrs2(1,2),'ko','MarkerSize',12,'LineWidth',2)
plot(ctrs2(2,1),ctrs2(2,2),'ko','MarkerSize',12,'LineWidth',2)
plot(ctrs2(3,3),ctrs2(3,4),'ko','MarkerSize',12,'LineWidth',2)
plot(ctrs2(4,3),ctrs2(4,4),'ko','MarkerSize',12,'LineWidth',2)
%
title('FKM New Clustering','FontName','Times New
Roman','fontsize',10,'fontweight','b');
legend('Cluster 1 new','Cluster 2 new','Cluster 3 new','Cluster 4
new','Centroids','Location','NW')
%
legend ('show');
set(legend,'Orientation','vertical','Location','SouthOutside',...
    'FontName','Times New Roman','fontsize',10,'fontweight','b');
%set(legend,'Location','SouthOutside');
hold off
groupmat=[repmat(1,mm1,1);repmat(2,mm2,1);repmat(3,mm3,1);repmat(4,mm
4,1)];
[k,l]=size(x2);
fold=4;
indices = crossvalind('Kfold',groupmat,fold);

for iii = 1:fold
test3(:,iii) = (indices == iii);
train3(:,iii) = ~test3 (:,iii);

Class(:,iii)=
knnclassify(x2(test3(:,iii),:),x2(train3(:,iii),:),groupmat(train3(:,
iii),:));
test_g(:,iii)=groupmat(test3(:,iii)~=0,1);

end
[mmm nnn] = size(Class);
for j=1:mmm
if Class(j)==test_g(j)
    result(j)=1;
elseif Class(j)~=test_g(j)
    result(j)=0;
end
end

total_percentage=mean(result)*100;
per_av=mean(mean(total_percentage,2));
```

AUTHOR’S PROFILE

Norizam Bin Sulaiman has completed his PhD in Electrical Engineering at the Faculty of Electrical Engineering, Universiti Teknologi MARA. He received his MSc. in Electronic Systems Design at the School of Electrical Engineering, Universiti Sains Malaysia, in 2005. He obtained his BSc. in Computer Engineering at Faculty of Electronics and Computer Engineering, Valparaiso University, Indiana, United States of America in 1995. Before taking his Master degree, he has been working as Quality Assurance (QA) Engineer at multinational company for 8 years. He works a lecturer in Faculty of Electrical and Electronics Engineering, Universiti Malaysia Pahang from 2005 until now.

LIST OF CONTRIBUTIONS

JOURNAL

[1] Norizam Sulaiman, Mohd Nasir Taib, Sahrim Lias, Zunairah Hj Murat, Siti Armiza Mohd Aris, and Noor Hayatee Abdul Hamid, “**Novel Methods for Stress Features Identification using EEG Signals**,” *International Journal of Simulation: Systems, Science and Technology (IJSSST)*, vol. 12, no. 1. pp. 27-33, February 2011.

[2] Norizam Sulaiman, Mohd Nasir Taib, Sahrim Lias, Zunairah Hj Murat, Mahfuzah Mustafa, Siti Armiza Mohd Aris, and Nazre Abdul Rashid, “**Electroencephalogram-Based Stress Index**,” *Journal of Medical Imaging and Health Informatics (JMIHI)*, vol. 2, no. 3, pp. 327-335, September 2012.

CONFERENCE PROCEEDINGS

[1] Norizam Sulaiman, Noor Hayatee Abdul Hamid, Zunairah Hj Murat, and Mohd Nasir Taib, “**Initial Investigation of Human Physical Stress Level Using Brainwaves**,” in *Proceedings of the 2009 Student Conference on Research and Development (SCOReD)*, 2009, pp. 1-4.

[2] Noor Hayatee Abdul Hamid, Norizam Sulaiman, Siti Armiza Mohd Aris, and Mohd Nasir Taib, "**Evaluation of Human Stress Using EEG Power Spectrum,**" in *Proceedings of the 6th International Colloquium on Signal Processing & Its Applications (CSPA)*, 2010, pp. 1-4.

[3] Siti Armiza Mohd Aris, Norizam Sulaiman, Noor Hayatee Abdul Hamid, and Mohd Nasir Taib, "**Initial Investigation on Alpha Asymmetry during Listening to Therapy Music**," in *Proceedings of the 6th International Colloquium on Signal Processing & Its Applications (CSPA)*, 2010, pp. 1-4.

[4] Sahrim Lias, Norizam Sulaiman, Zunairah Hj Murat, and Mohd Nasir Taib, "**IQ Index using Alpha-Beta Correlation of EEG Power Spectrum Density (PSD)**," in *Proceedings of the* IEEE *Symposium Industrial Electronics and Applications*, 2010, pp. 612-616.

[5] Norizam Sulaiman, Mohd Nasir Taib, Siti Armiza Mohd Aris, Noor Hayatee Abdul Hamid, Sahrim Lias, and Zunairah Haji Murat, "**Stress features identification from EEG signals using EEG Asymmetry and Spectral Centroids Techniques**," in *Proceedings of the IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, 2010, pp. 417-421.

[6] Norizam Sulaiman, Mohd Nasir Taib, Sahrim Lias, Zunairah Hj Murat, Siti Armiza Mohd. Aris, and Noor Hayatee Abdul Hamid, "**EEG-based Stress Features Using Spectral Centroids Techniques and k-Nearest Neighbor Classifier**," in *Proceedings of the UKSim 13rd International Conference on Modeling & Simulation*, 2011, pp. 69-74.

[7] Siti Armiza Mohd Aris, Mohd Nasir Taib, Sahrim Lias, and Norizam Sulaiman, "**Feature Extraction of EEG Signals and Classification using FCM**," in *Proceedings of the UKSim 13th International Conference* on *Modeling and Simulation*, 2011, pp. 54-58.

[8] Norizam Sulaiman, Mohd Nasir Taib, Sahrim Lias, Zunairah Hj Murat, Siti Armiza, Mohd Mahfuzah Mustafa, and Noor Hayatee Abdul Hamid, "**Intelligent System for Assessing Human Stress Using EEG Signals and**

Psychoanalysis Tests," in *Proceedings of the* 3rd *International Conference on Computational Intelligence, Communication Systems and Networks*, 2011, pp. 363-367.

[9] Norizam Sulaiman**,** Mohd Nasir Taib, Sahrim Lias, Zunairah Hj Murat, Siti Armiza Mohd Aris, Mahfuzah Mustafa, and Nazre Abdul Rashid, "**Development of EEG-Based Stress Index**," in *Proceedings of the 2012 International Conference on Biomedical (ICoBE)*, 2012**,** pp. 461-466.

LIST OF PATENT

[1] A Method for Evaluating Human Brain Attribute, **UI 20110031**.

LIST OF AWARDS

[1] 7TH Invention, Innovation & Design, IID2009, UiTM, 12-14 January, 2010. Project title: "**Portable Brainwave Balancing Index using EEG**" - Gold Medal.

[2] Malaysia Technology Expo, MTE 2010, PWTC, 10-13 February, 2010. Project title: "**Portable Brainwave Balancing Index using EEG**" - Gold Medal

[3] 8TH Invention, Innovation & Design (Special Edition), IID2010SE, 12-14 October, 2010. Project title: "**Intelligent System for Human Brainwave Centric Evaluation**" - Gold Medal

[4] Malaysia Technology Expo, MTE 2011, KLCC, 17-19 February, 2011, Project title: "**Intelligent System for Human Brainwave Centric Evaluation**" – Silver Medal.