EEG-based Mental Workload and Emotion Detection using Machine Learning and Deep Learning

Major 2 Project Dissertation

Submitted in partial fulfillment of
the requirements for the award of the degree
of

Master of Technology
in
Department of Computer Science and Engineering
by

Debarshi Nath
(Roll No: 2K18/CSE/04)

under the guidance of

Dr. Divyashikha Sethia
Assistant Professor
Department of Computer Science and Engineering

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
DELHI TECHNOLOGICAL UNIVERSITY, DELHI
DECEMBER, 2019

DEPARTMENT OF COMPUTER SCIENCE ENGINEERING
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi - 110042

DECLARATION

I, Debarshi Nath, Roll No. 2K18/CSE/04, student of M.Tech (Computer Science and Engineering), hereby declare that the project dissertation titled **EEG-based Mental Workload and Emotion Detection using Machine Learning and Deep Learning** which is submitted by me to the Department of Computer Science and Engineering, Delhi Technological University, Delhi in partial fulfilment of the requirement for the award of the degree of Master of Technology, is original and not copied from any source without proper citation. This work has not been submitted anywhere for the award of degree, diploma, fellowship or other similar title or recognition to the best of my knowledge.

Debarshi Nath

Place: Delhi
Date: Aug 31, 2020

Debarshi Nath
Roll No. 2K18/CSE/04
Department of Computer Science and Engineering,
Delhi Technological University

DEPARTMENT OF COMPUTER SCIENCE ENGINEERING
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi - 110042

CERTIFICATE

I certify that the Project Dissertation titled **EEG-based Mental Workload and Emotion Detection using Machine Learning and Deep Learning** which is submitted by Debarshi Nath (Roll number: 2K18/CSE/04), Department of Computer Science and Engineering, Delhi Technological University, Delhi in partial fulfilment of the requirement for the award of the degree of Master of Technology, is a record of the project work carried out by the student under my supervision. To the best of my knowledge this work has not been submitted in part or full for any degree or diploma to this university or elsewhere.

Divya Sethia

Place: Delhi
Date: June 30, 2020

Dr. Divyashikha Sethia
Assistant Professor,
Department of Computer
Science and Engineering,
Delhi Technological
University

ACKNOWLEDGEMENT

The success of a project requires help and contribution from numerous individuals and the organization. Writing the thesis of this project work gives me an opportunity to express my gratitude to everyone who has helped in shaping up the outcome of the project. I express my heartfelt gratitude to my project supervisor **Dr. Divyashikha Sethia**, Department of CSE, giving me the opportunity to do my project work under her guidance. Her expertise and constant backing have pushed this project to this extent even at a trying time like this. I extend my gratitude to **Dr. S. Indu**, Department of ECE for helping this project with her knowledge and motivation. I also express my gratitute to **Dr. Suman Kushwaha**, HOD, Department of Neurology, Institute of Human Behaviour and Allied Sciences (IHBAS), Delhi for her expert knowledge and guidance in understanding EEG and physiological parameters and their clinical aspects. Their constant support and encouragement has made me realize that it is the process of learning which weighs more than the end result.

I express my gratitude to my labmates, for without their contribution this project would not have been possible. I thank **Diksha Kalra** for helping this project with her immense knowledge of signal processing. I reveal my thanks to **Anubhav** and **Mrigank Singh** for their expertise in machine learning and their strong determination in ensuring that the project goes on despite the bumps on the road.

I also reveal my thanks to all my friends and my family for consistent support.

Debarshi Nath

Place: Delhi
Date: Aug 31, 2020

Debarshi Nath
Roll No. 2K18/CSE/04
Department of Computer Science and Engineering,
Delhi Technological University

ABSTRACT

Mental stress has significant impact on critical thinking, problem solving, behaviour, social interaction, and general intelligence. Electroencephalography (EEG) is a simple method which gives an idea about the potential generated on the surface of the brain which helps in understanding the functionality of the brain. EEG finds its use in various bio-medical and bio-informatics research works where the EEG patterns can be used to predict and analyse a person's mental state and awareness.

In this work, we propose an effective approach to detect mental stress. We address two problems in EEG-studies that have generated considerable interest in recent times: (a) Emotion Recognition, and (b) Mental Stress Detection. We make use of publicly available benchmark EEG datasets to investigate these problems and unravel key information from these studies. From the Emotion Recognition study, we formulate an effective emotion recognition approach and highlight the differences between subject-dependent and subject-independent models. We achieve the best average classification accuracy of 93.91%. The stress detection study on publicly available dataset 'EEG During Mental Arithmetic Tasks' finds the theta and alpha power of EEG signals most indicative of cognitive workload. The study also highlights the significance of frontal EEG channels in mental stress studies. We achieve the best accuracy of 93.05% for this study. We apply the findings of these studies on the EEG recordings that we gather from students of Delhi Technological University. We obtain the best classification accuracy of 90% for this study. We propose an efficient EEG-based stress detection system that can be used to determine stress in students in real world.

Contents

List of Figures

List of Tables

List of Abbreviations

ANOVA	Analysis of Variance
CNN	Convolutional Neural Network
Cog-MAPS	Cognitive-Mental Assessment using Physiological Signals
DBN	Deep Belief Networks
DEAP	Database for Emotional Analysis using Physiological Signals
EEG	Electroencephalogram
KNN	K-Nearest Neighbor
LSTM	Long Short-Term Memory
MIST	Montreal Imaging Stress Task
MLP	Multi Layer Perceptron
PSD	Power Spectral Density
RFECV	Recursive Feature Elimination using Cross-Validation
RNN	Recursive Neural Network
SAE	Stacked Auto-Encoder
SVM	Support Vector Machine

Chapter 1

Introduction

The human brain consists of billion of cells, a large number of which are neurons, and others which aid and facilitate the activity of neurons. Whenever any activity occurs it generates electrical impulse in the brain due to which thousands of neurons fire in sync. Emotions and responses are determined by the communication between neurons within our cranium. Brainwaves are generated by electrical pulses fired in sync from masses of neurons communicating with each other. Brainwaves are detected using sensors placed on the scalp. Electroencephalogram is the test that is used to detect these brainwaves. These brainwaves can be decomposed into bandwidths of frequencies according to their functions, but are commonly recognized of as a continuous flow of consciousness: from slow, loud and functional - to fast, subtle, and complex.

1.1 EEG applications in recent years

Electroencephalography (EEG) in recent years has become a significant tool with its applications and utility in neuroscience, Brain–Computer Interfaces (BCI's), and commercial applications as well. Analytical tools prevalent in EEG studies frequently use machine learning techniques to uncover relevant information for neural classification and neuroimaging.

Machine learning is an application of Artificial Intelligence (AI) that provides systems the ability to learn patterns on its own and improve with time and data without being explicitly programmed. Machine learning focuses on the development of algorithms and programs that can access data and use it to learn for themselves. Deep Learning is a subfield of machine learning concerned with algorithms inspired by the structure and function of the brain called Artificial Neural Networks (ANN).

Evidence has been found that EEG characteristics can be used as an indicator (a biomarker) of some diseases. In a project funded by The Michael J. Fox Foundation, researchers have found that there are significant differences in the EEG data of different Rapid-Eye Movement Behaviour Disorder (RBD) patients compared to healthy people [5].

Justin Alvey [6] and Thomas Fulvio [7] in their 2019 articles, describe how to use these EEG signals and deep learning to classify sub-vocalized words — specifically by reading the electrical brain activity using an EEG/EMG sensor, setting up module for processing and acquiring labelled training data, and creating a custom Convolutional Neural Network (CNN) for classification. Using Machine Learning to categorise EEG signals from the cranium and transform them to words, the desired goal for Fulvio would be an app that is reading the thoughts of a person who is trying to speak and output those words to those they are trying to communicate with. This can be of great benefit and significance to people who are senile and bed-ridden and unable to speak up due to disorders.

One application of machine learning and deep learning in the literature of EEG-based tasks is the problem of studying stress and attention. A number of studies have come up in recent years where artificial intelligence has performed significantly well in classification tasks. Studies have employed not only EEG gathered from subjects undergoing these tests but also a number of physiological parameters and biomarkers which have been designated as implication of stress and attention levels.

Hans Selye [8] defines stress as an unspecific response of the human body to the demand of a task such as mental, emotional, or physical. Although often clubbed together, and sometimes even used interchangeably, there are major differences between stress and cognition. Breed et al. [9] define cognition as the ability of the thought process to bring together facts and imagination to find a solution. This cognitive ability to interpret, understand, plan and have foresight is what separates human beings from other animals. Mental activities and exercises are also a test of the cognitive ability of a person. The relation between cognition and stress, with stress being a consequence of cognitive workload and the effects of stressful stimuli on perception have been documented in literature [10, 11]. Thus changing the instructions and demands of a task, required cognitive adaptions and performances have a taxing effect on the mind of an individual and such mental tasks have been recognised as stressors in psycho-physiological research [12]. Various cognitive-workload studies utilise these 'stressors' to induce short-term stress on the participants.

Stress has been a rampant threat in the society in recent times and there is a scope and necessity of studying it and tackle the negatives that accompany stress and anxiety. This has been a motivation for this study.

1.2 Problem Statement

In literature, a number of studies have come up to learn about stress and attention. A few of this studies have employed machine learning techniques and deep learning. Apart from studying the EEG of the subjects, some of these studies have employed the performance of physiological parameters in classification of the stress states [13]. This study aims to address the following issues related to the study of stress and attention:

- The input formulations to machine learning and deep learning models that can be used for EEG-based stress detection.
- Classification architectures that can be efficient to for this study.
- The performance of EEG in Emotion Recognition tasks.
- The relevant features that can be used to build an effective stress detection model.
- Finding out the EEG channels that are most significant for detecting stress.
- Proposing an effective stress detection model that can be used to detect mental stress using feedback from physiological sensors.

The remainder of this thesis is divided into six chapters. In Chaper 2, we describe a few basic terminologies and EEG classification tasks. In Chapter 3, we report and analyse recent publications and literature relevant to our study. In Chapter 4, We report the performance of machine learning and deep learning architecture in a popular EEG-based problem- emotion recognition. Chapter 5 reports our extensive analyses on an existing EEG dataset, aimed at detecting cognitive workload in individuals undergoing a mental arithmetic task. Chapter 6 contains the report of our stress detection study on 5 subjects from Delhi Technological University. Finally in Chapter 7, we include our concluding remarks on the work and highlight the scope of future work

Chapter 2

Literature Review

2.1 Brainwaves

The brainwaves emanating from the human cranium can be decomposed into bandwidths of frequencies according to their functions, but are commonly recognized of as a continuous flow of consciousness: from slow, loud and functional - to fast, subtle, and complex. These bands of frequencies are extremely sensitive, highly informative and vary drastically across different EEG tasks. For these reasons, these frequency bands have been extensively studied and explored since decades. These bands have been described in brief below [14].

1. **Delta Waves (0.5-4 Hz)**: It assesses the depth of sleep. The stronger the delta rhythm, the deeper the sleep. They are often found in infants and young children. Adequate generation of delta waves makes us rejuvenated after sleep. They are generated during deep meditational states and in state of dreamless sleep. Delta waves take away external awareness and are source of empathy. Too much presence of delta waves has been linked to brain injuries, learning problems and inability to think. Inadequate delta waves in brainwaves are signs of inability of the body to rejuvenate and revitalize and indication of poor sleep.

2. **Theta Waves (4-8 Hz)**: Theta brainwaves are emitted mostly in sleep but are also a major brainwave component in deep meditation. Theta can be seen as a portal to learning, instinct and memory. Excessive theta activity has been linked to bouts of depression. It also might be a sign of hyperactivity, impulsivity, inattentiveness. Inadequate theta waves are sign of anxiety, poor emotional awareness and stress. It appears in the beginning stages of sleep (drowsiness). It is also associated with a myriad of cognitive processing events such as memory encoding in case of difficult tasks.

3. **Alpha Waves (8-13 Hz)**: Alpha brainwaves are significant during dormant but continuous flow of thoughts, and in a few meditative states. Mental coordination, alertness, body/mind integration and learning are all dependent on alpha activity in the brain. It calms us down when required and relaxes the mind. If we undergo mental stress, a phenomenon called "alpha blocking" may occur which is characterized by very little alpha activity and excessive beta activity. Alpha wave activity is enhanced in relaxed state and also alpha power is more in case of happy emotion as compared to when a person is sad. Alpha waves have been considered to have an

overlapping between casual mindfulness and furthermore absent-mindedness. Too much alpha waves has been linked to daydreaming, inability to focus and high relaxation. Lack of alpha waves is associated with stress, anxiety and insomnia. Consumption of alcohol and some antidepressants show increase in alpha waves.

4. **Beta Waves (13-30 Hz)**: Beta brainwaves dominate our general conscious state of mind when attention is concentrated towards cognitive activities of the external world. Beta exists when we are alert, active, attentive, engaged in solving problems, judgement and decision making, or focused mind activities. Having optimum amount of beta waves makes us focussed and perform school or work-related tasks with ease. But people with too much beta may experience excessive stress and/or anxiety. Presence of high amounts of beta waves is linked to excessive anxiety, adrenaline rush, high arousal, inability to relax and sleep and stress in human system. Naturally low beta presence is found in case of daydreaming, depression and poor cognition. Consumption of coffee, energy drinks and various stimulants increase beta activity in brain.

5. **Gamma Waves (30-60 Hz)**: Gamma brainwaves are the fastest of the brainwave spectrum, and are characteristics of simultaneous information processing from different brain areas. It dominates when a person tries to combine two different senses such as sound and sight. These are observed in high processing activities as well as cognitive functioning. It has been established that mentally challenged individuals with learning disabilities tend to have lower gamma activity than that of average individuals.

2.2 Classification Problems in EEG-based Tasks

With so many applications of machine learning and deep learning in the field of EEG based problems having come up in recent years, some natural questions come up:

- What activities can be done with the application of Machine Learning and Deep Learning using EEG?
- For deep learning, what EEG recognition tasks have been explored?
- What kinds of features have been effective in this field of problems?
- Are specific deep learning models more effective in solving specific EEG – based problem?

Very recently, Craik et al. [1] and Roy et al. [15] have tried to address these questions by studying and systematically categorizing the EEG classification tasks and their deep learning approaches in an effort for easier understanding of the current prevalent research in this field. Their findings have been discussed in the next sections.

2.2.1 Categorizing the EEG Tasks

The tasks presented within the studies of Craik et al. [4] fell into six major groups: emotion recognition, motor imagery, mental workload, seizure detection, sleep stage scoring, and event related potential detection. Apart from these groups, there were some other studies- which include Alzheimer's disease classification, depression studies, etc.

1. **Emotion recognition tasks**: Emotion recognition tasks characteristically involve subjects being shown video clips or listening audio clips which have been assigned specific emotion and rated by experts prior to the showing, like that in the creation of DEAP dataset [16]. EEG was collected during these viewings and was followed by an emotion self-assessment. The self-assessment and original emotion categories and ratings was then converted into recognized emotion dimensions or parameters, namely valence, arousal, dominance and liking values, used in a widely used system to describe emotions. The chief goal of emotion recognition and classification studies, in a general sense, is to help us better understand the emotional state of the individual at that time.

2. **Motor imagery tasks**: Motor imagery tasks have experiments which involve the subject envisage certain muscle movements on muscles of limbs and/or the tongue. Their applications are mostly Brain-Machine Interface related, where BMI applications are eventually expected to classify a subject's intended motion or movement.

3. **Mental workload tasks**: Mental workload tasks typically involve measuring and collecting EEG signals from an individual's brain periphery while the subject was in varying degrees of complexity of the mental task. There are two general areas where this type of task can be applied: cognitive stress monitoring or BMI performance monitoring. An example of such a task is the one involved in the dataset 'EEG During Mental Arithmetic Tasks' [17]. The participants performed a mental arithmetic task during the course of which their EEG recordings were captured.

4. **Seizure detection tasks**: For seizure detection studies, EEG signals are recorded in patients with epilepsy during seizure periods and post-event periods. Non-epileptic patients have registered EEG signals as a control group for some datasets, so that a comparison is available between a sick person and a fit person. Such experiments were planned for the potential application for the prediction of possible seizures and for the pre-emptive warning of epileptic patients.

5. **Sleep stage scoring tasks**: Sleep stage score studies have the least number of studies, with experiments that record subjects ' EEG signals overnight. These signals were scored by experts and categorized into the recognized sleep stages 1, 2, 3, 4 and the rapid eye movement phases. The research objective of this type of study is to reduce the dependency on trained experts for the analysis and understanding of a person's sleep.

6. **Event related potential tasks**: Studies that focused on detecting and classifying potential associated with events typically record EEG from subjects undergoing a visual presentation task. In these tasks, the subject monitors a rapid sequence of images or letters with a focus on specific indicators. Once a specific letter or image appears, a stereotypical response is shown in the EEG data.

2.2.2 Input formulation

It has been generally found that the input to a classification model related to application in EEG – based tasks to fall in 3 categories:

1. Signal Values

2. Calculated Features

3. Images

The selection of these features relies, to a great extent, on the choice of classifier used. Signal values are the category of input formulation that includes using the raw or averaged signal for input into the classifier. Complex Value Transformation (CVT) is another input formulation that fall into this category. Neural networks have shown promise when signal values have been used for classification tasks, encouraging an end-to-end learning philosophy. These techniques have encouraged studies which involve feeding raw EEG signal values directly into the neural network without processed features, which may signify the possibility of directly studying raw EEG data with deep learning for faster and efficient real-time output.

Calculated features have often formed an important and major group of input for the classifiers studied in the survey of Craik et al. [1], accounting to about 41% of these studies. EEG data frequently analyzed in frequency domain. Power spectral density (PSD), wavelet decomposition, and statistical measures of the signal like mean, variance and standard deviation are three of the most popular input formulations used in the studies of the review.

A number of neural networks, primarily CNN's, use spectrograms generated from the EEG data as input feature. Spectrograms have long been used as a postprocessing tool to visualize the data. However, CNN's capability to learn from images enables spectrograms to be used an input formulation into the classifier. Examples of a few other image input formulations include Fourier feature maps and 2D/3D grids.

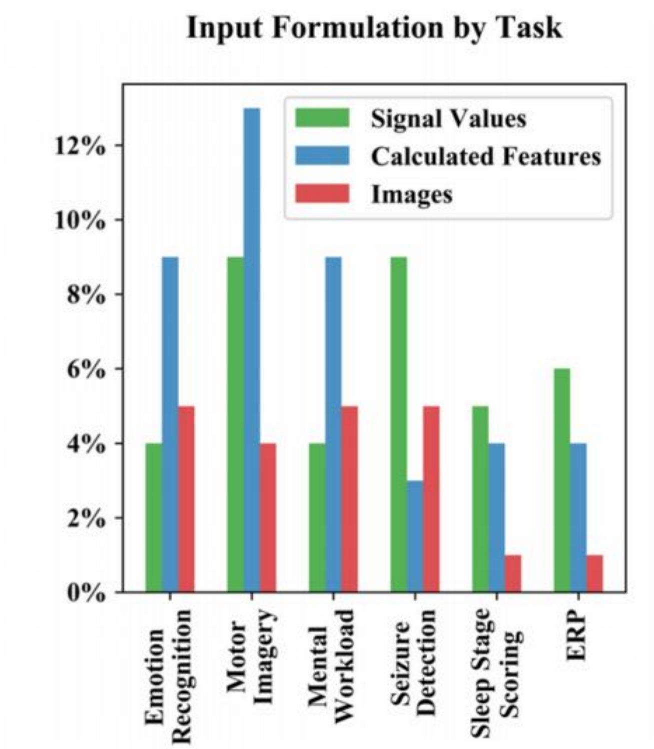


Figure 2.1: The figure represents the percentage of input features used in each task [1]

2.2.3 Architecture design choices

The choice of input formulation for the EEG-based tasks relied heavily on the task and deep learning architecture. Craik et al. [1] found 6 basic architectures used in these tasks: Convolutional Neural Networks (CNN), Deep Belief Networks (DBN), Recurrent Neural Networks (RNN), Multi Layer Perceptron (MLP), and Sparse Autoencoder (SAE). The inputs going into these architectures are shown in the diagram below:

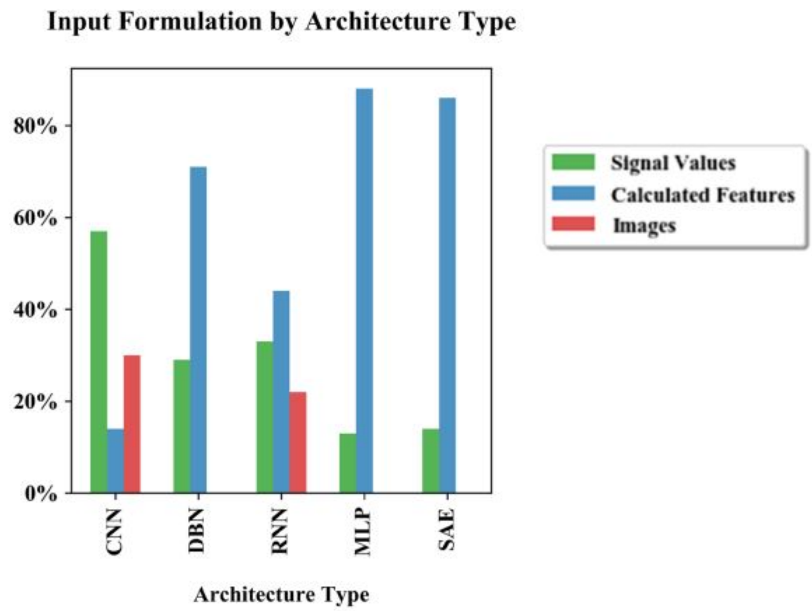


Figure 2.2: The inputs to different architectures used in EEG tasks [1]

From figure 2.2, while we see no particular preference that address all the tasks, works that used either MLP or SAE showed an distinct affinity towards calculated features as input. Both CNN and RNN studies saw instances where all three input formulations have been used in literature. The choice of input for DBN studies were divided between signal values and calculated features, with calculated features being the most preferred. Images went as features only into CNN and RNN architectures. There were no studies in which the accuracies of using signal values and calculated features are contrasted, suggesting room for more study.

Emotion recognition, motor imagery, and sleep stage scoring tasks do not seem to have a particular affinity towards choice of specific deep learning models (interpreting from figure 2.3. Seizure detection studies were seen to be using either CNN's or RNN's, with the major percentage of studies using RNN's in these tasks as compared to other tasks. Sleep stage scoring tasks had the highest percentage of studies using hybrid architecture formulations.

Mental Workload detection studies too didn't have a particular affinity towards single architecture. Further, an efficient deep learning architecture requires a sufficiently large database for training the classification model. For these reasons, machine learning has gained prominence in the recent EEG studies. As machine learning has gained momentum for smaller datasets, research has also been extensively done towards finding out features that can be used as for task-specific research. For the development of an effective model of classification, our research too has been directed towards finding out the features that can leverage the EEG signals towards development of an effective model that can work

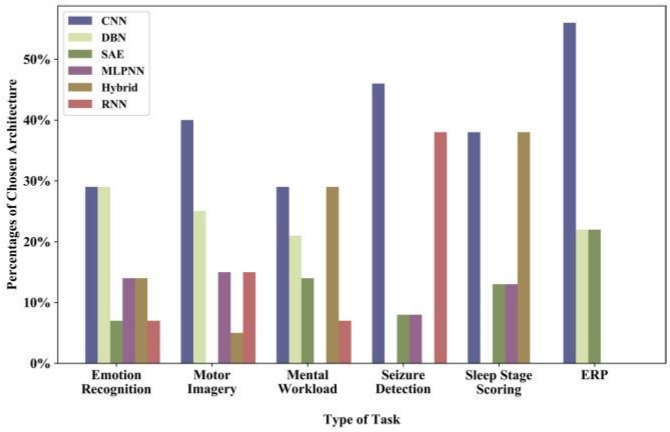


Figure 2.3: The trend of use of specific architectures in specific EEG-related tasks [1]

not only with benchmark EEG stress detection datasets, but also for a real-time stress recognition model using EEG signals and other physiological biomarkers as feedback from portable sensors in long term.

2.3 Features

Frequency domain features of EEG are of particular interest in EEG Tasks. From our subsequent work in the emotion recognition (discussed in detail in Chapter 4), we found its efficiency in classification. We were now left with finding suitable frequency domain features for the task of stress detection. We identified and selected four frequency domain features (The classification using these features is discussed in detail in Chapter 5).

A great number of frequency domain features are derivative of the Power Spectral Density (PSD). It is a measure of power contained in a signal at each point of a signal plotted against frequency. A typical PSD plot is shown in figure 2.4.

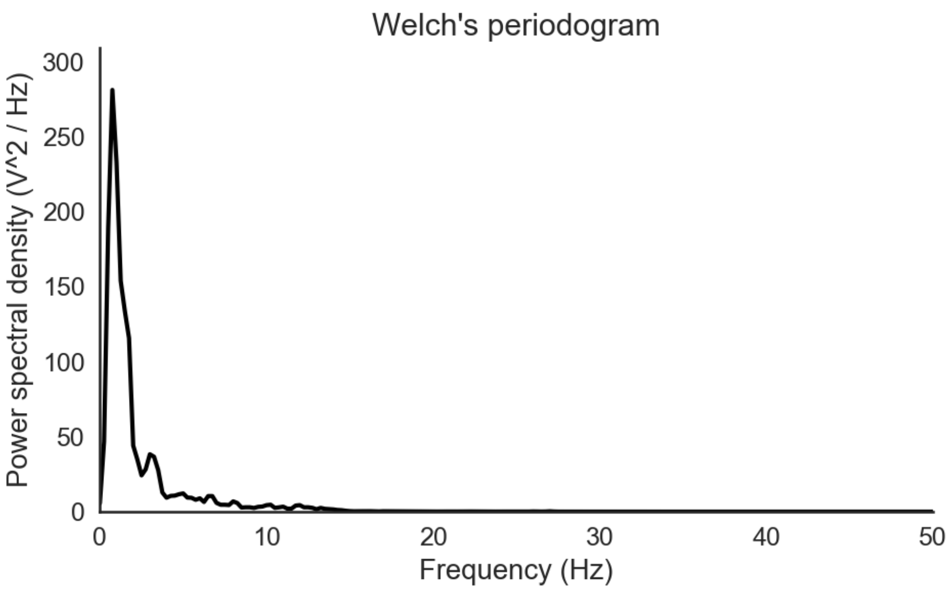


Figure 2.4: Power Spectral Density of an EEG channel [2]

We implement the Welch method [18] to determine the Power Spectral Density (PSD) of the EEG signals. The four PSD-derivative features that we use in stress detection (discussed in more details in Chapter 5) are described below.

2.3.1 Spectral Entropy

The work of Tian et al. [19] has shown promise for Spectral Entropy to be used as a feature for mental workload studies. Spectral Entropy is defined as the Shannon Entropy of the normalised Power Spectral Density (PSD) of a signal. The Spectral Entropy of a signal x is given as:

$$SE = -\sum_{f=0.5}^{30} P(f) log[P(f)] \tag{2.1}$$

where $P(f)$ is the normalised Power Spectral Density (PSD) of the signal.

2.3.2 Relative Power

Using the PSD, we also calculate the Absolute Power for each frequency band from Delta to Beta (also known as bandpower of the particular band). Various studies utilise Relative Power for EEG based mental stress and cognition detection [20–22]. Thus, we compute Relative Power from the Absolute Power of bands using the Equation 2.2.

$$Relative\ Power = \frac{Absolute\ Power\ of\ band}{Total\ Power} \tag{2.2}$$

2.3.3 Alpha-Beta Ratio

An earlier study reports a decrease in Alpha activity and a parallel increase in Beta activity during stressful mental arithmetic task [23]. This study motivates us to examine the activities of these two bands with the Alpha-Beta ratio, Equation 2.3.

$$Alpha - Beta\ Ratio = \frac{Absolute\ Power\ of\ Alpha\ band}{Absolute\ Power\ of\ Beta\ band} \tag{2.3}$$

2.3.4 Theta-Alpha Ratio

Some studies [24, 25] report an increase in the Theta activity with increase in mental workload. Thus, we expect the Theta-Alpha ratio to increase during the mental activity phase. Similar effect is also observed by [26,27]. Equation 2.4, describes the Theta-Alpha ratio as :

$$Theta - Alpha\ Ratio = \frac{Absolute\ Power\ of\ Theta\ band}{Absolute\ Power\ of\ Alpha\ band} \tag{2.4}$$

We used these features to design our classification model for stress detection studies. The details of these studies are discussed in detail in Chapter 5 and Chapter 6.

Chapter 3

Related Works

3.1 Recent Works in EEG-based Stress Detection

During the course of this project, we studied several papers, with the goal of identifying recent papers that can contribute towards the consequent assessment and development of a stress detection model using EEG. Out of those several papers, we identified 42 papers that have recently been published in the domain of Mental Workload detection. The search was made using Google Search Engine using the keywords like 'Stress Detection', 'Stress Recognition', 'Mental Workload detection', 'EEG', 'Electroencephalography', 'Physiological Signals', 'Machine Learning', 'Deep Learning'. The papers have been identified and filtered pertaining to the following criteria:

- **Electroencephalography**: Only the studies that employ EEG signals for prediction and analyses have been included for our study. Multimodal studies have been included, but only the ones that also include the results using lone EEG signals.
- **Machine Learning and Deep Learning**: The studies must employ classic machine learning techniques or deep learning architectures.
- **Task**: Only the studies that deal with short-term mental stress or cognitive workload detection are considered.
- **Time**: Due to the large volumes of studies published in this domain so far, we have only considered the papers that have published in the year 2011 and after.

Out of these 42 papers, the prominent ones are mentioned in brief below.

Hafeez et al. [28] used a 50Hz notch filter for removal of power line noise, and applied ICA, and filtering techniques for removal of ocular and muscular movement noise. They evaluate Frequency domain features such as Power Spectral Density of Alpha, Beta, and Theta bands for 14 subjects. The study concludes with a remark of 85% of students experiencing stress before the examination.

Sharma and Chopra [29] employed Instantaneous Frequency of various Intrinsic Mode Functions (IMFs) obtained by applying Hilbert Huang Transform (HHT) as a Time-Frequency domain feature. The data was cleaned using wavelet decomposition and a joint combination of 0.75Hz high-pass filter and 45Hz FIR filter. This work reports a maximum accuracy of 92.86% using the Support Vector Machine (SVM) classifier.

Blanco et al. [30] propose a real-time stress prediction system using the Emotiv Epoc Headset on 18 subjects. The raw EEG data obtained from headset was cleaned by subtracting the least-squares line of best fit, and then the data was passed through a bandpass filter network of Chebyshev type II filters. Band Power and Root Mean Square (RMS) value of signal were extracted as a feature from the artefact free data, and Logistic Regression (LR), QDA and K-NN were used for prediction achieving a maximum accuracy of 78.70%.

Hasan and Kim [31] employed the publically available DEAP dataset comprising of 32 subjects for stress prediction. The data was averaged out to standard reference, and bandpass filter of range 4-45Hz was employed for data cleaning. Various Time-domain features like RMS, Peak-to-Peak value, Kurtosis, Skewness, Hjorth parameters like Mobility and Complexity along with Time-Frequency domain features using Wavelet transform was computed. Time domain and Time-Frequency domain features were combined, and feature selection method Boruta was applied, achieving a classification accuracy of 73.38% for KNN classifier.

Nagar and Sethia [32] proposed a real-time stress prediction system, using a single electrode Neurosky Mindwave device and collected data of 63 students. Thresholding technique was employed to remove EEG data having an amplitude above 100uV, and frequency components above 50Hz were removed using a suitable low-pass filter. The band power ratio of different EEG bands, namely Alpha, Beta, Delta, and Theta, was computed as Frequency domain features. PSS-14 questionnaire response of students along with extracted features was given as input to the KNN classifier, achieving a maximum accuracy of 74.43%.

Vourkas et al. [33] used a 16 channel EEG device to gather the data of 20 participants. A 40 Hz low-pass filter was used to remove the higher frequencies, and Hjorth parameters were computed as features. Testing accuracy of 83% was achieved using ANN classifier for discrimination of EEG signals during mental task and rest phases.

Islam et al. [34] employed the dataset of 15 participants which was collected using the BIOPAC data acquisition unit MP36R. FIR bandpass filter and Hanning window were used for removal of noise and signal smoothing. Different statistical features such as mean value, standard deviation, skewness, kurtosis, maximum value were computed to train SVM, KNN and ANN classifiers on the EEG database for the cognitive tasks. They report the classification accuracy 95.21%, 90.88%, and 94.39% for SVM, KNN and ANN respectively.

Türk et al. [35] use Hjorth parameters to extract features from EEG signals for numerous mental tasks. Using Artificial Neural Networks (ANN), they report 77.61% average classification accuracy for all sets of mental tasks.

Yazıcı and Ulutaş [36] evaluate Minimum, Maximum, Average, Standard Deviation, and other Time-domain features to train the SVM classifier. They report 91.13% classification accuracy emulating the winning accuracy of 88.89%.

Oh et al. [37] employs Hjorth parameters to improve the classification accuracy in motor imagery case application of EEG by 4.4%. In [38–40] authors, use Hjorth parameters with other Time-domain features for different applications of EEG signals.

Ahammed and Ahmed [41] used the publicly available EEG During Mental Arithmetic Task [17] database for stress prediction during rest and mental task. Multivariate Multiscale Entropy (MMSE) was used as a feature, and the channel selection was applied to determine stressed regions of the brain. The SVM classifier was employed achieving 90% accuracy for rest v/s mental stress and 87.5% accuracy for good v/s bad counting.

Table 3.1: Comparison of best performance with some recent and prominent works in stress detection

Paper	Dataset/Task	Feature Domain	Classifier	Accuracy
Sharma et al. [29]	EEG recorded during aptitude test	Time-Frequency	SVM	92.86%
Blanco et al. [30]	Stroop Test	Frequency	KNN	78.70%
Nagar et al. [32]	Perceived stress in students	Frequency	KNN	74.43%
Ahammed et al. [41]	EEG During Mental Arithmetic Tasks	Time	SVM	90%
Priya et al. [42]	EEG During Mental Arithmetic Tasks	Frequency	KNN	99.42%
Our work	**EEG During Mental Arithmetic Tasks**	**Time** **Frequency** **Time-Frequency**	**SVM**	**97.32%** **97.5%** **98.61%**

3.2 Findings

From the literature survey, we were provided with multiple insights into the recent trends and updates in the domain of stress detection and classification. Of these insights, we present a few briefly in the following section. These findings are aimed at getting more information of the classifiers, features and filtering techniques that have been able to perform with significant accuracies.

3.2.1 Choice of Classifier

The first task was to find the popular choice of classifiers among the voluminous list of classifiers that have been used in the classification tasks for mental workload detection.

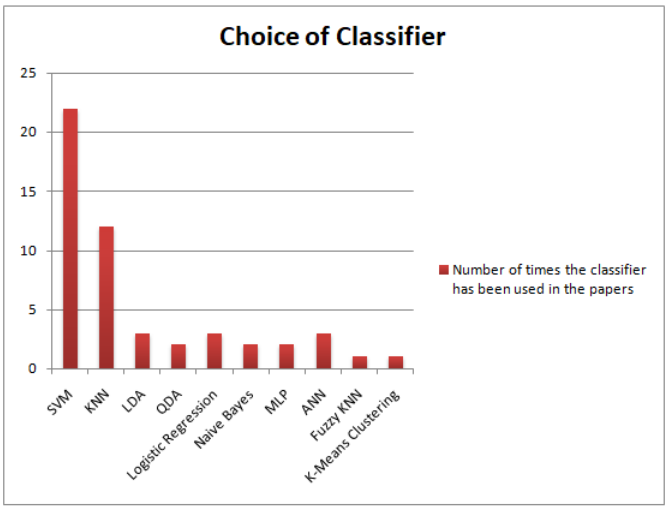


Figure 3.1: The choice of classifier in the studies

Among the classifiers used in the 42 papers, Support Vector Machine (SVM) was the most popular choice of classifier. It was used in 22 out of 42 papers, making it an obvious classifier to explore for performance on an EEG dataset. K-Nearest Neighbors (KNN) followed it with 12 instances of usage in the papers. The usages of all the classifiers is shown in figure 3.1.

3.2.2 Performance of Classifiers

The performance of the classifiers on EEG signals for classification is probably an even more interesting and important avenue to look into. This will give us more insight into the performances so far, and narrow down our search for a classifier that will perform well on our dataset.

In the studies, plenty of works achieved an accuracy of above 90% for 2-class stress classification. Hence in our representation of such works in 3.2, we only visualize the classifiers with classification accuracy of 90% and above. Since we are working towards development of a model for binary stress classification, these accuracies are reported only for 2-class stress detection and classification.

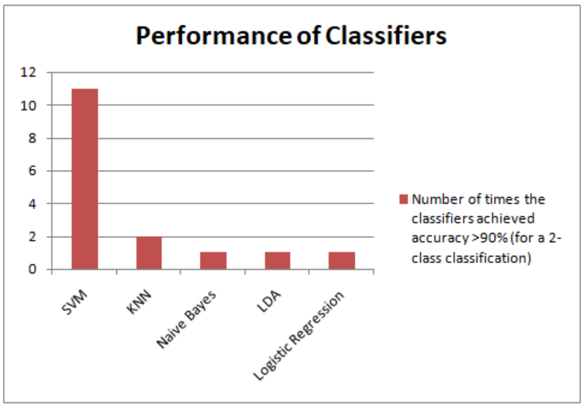


Figure 3.2: Performance of Classifiers in a two-class classification

From figure 3.2, we see that in this category too, SVM was the top performer, achieving a classification accuracy of 90% and above in 11 papers. KNN achieved the benchmark performance twice, while each of the other classifiers achieved that performance once.

3.2.3 Choice of Features

A requirement for machine learning models is that it requires explicit features as input for an effective classification. While this search for appropriate feature is often a vague venture in presence of large amount of unstructured data, finding appropriate features and consequent selection of the important ones can drastically improve the performance of the classifiers both in terms of speed as well as accuracy. In absence of big EEG datasets, the search for the right feature for the right mental task has been looked into quite often.

The EEG features can be conveniently divided broadly into 3 categories based on the domain of analysis (i.e., expressed as a function of the concerned domain). These are-

1. **Time domain**: Statistical features like mean, variance, maximum of the amplitudes, RMS value, Shannon Entropy, etc. which can be generated from a time series.

2. **Frequency domain**: Features generated after transform of the signal from time domain to frequency domain, like PSD, absolute power, power ratios, Spectral Entropy, etc.

3. **Time-Frequency domain**: These features are often generated by transforming the time series signal into frequency series and then performing a statistical value from the series.

Often categorized and studied separately, these features are of great significance in EEG classification tasks. We try to explore their usage in our papers at hand. The choice of features is shown in figure 3.3.

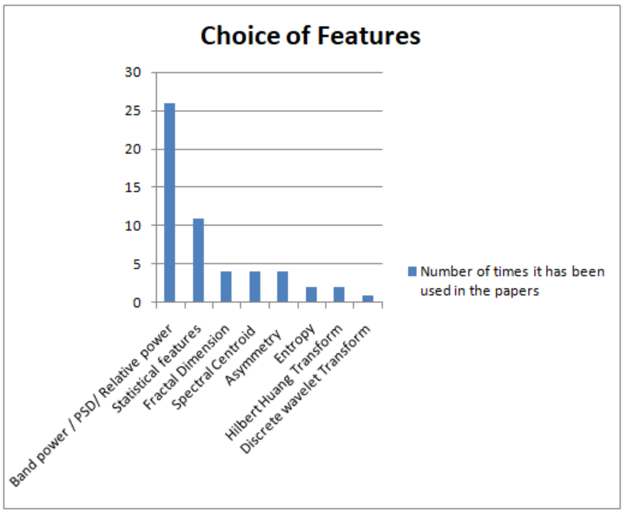


Figure 3.3: Choice of Input Features

Bandpower and Power Spectral Density (PSD) based features have been explored the maximum for a total of 26 times in the papers. These are frequency domain features. Statistical features like Mean, Variance, Root Mean Square (RMS) Value have been explored in 11 papers. We see a glaring affinity towards frequency domain features.

3.2.4 Performance of Features

While the choice is explored, the performance of these features, obviously, cannot be ignored. We try to find their impact in classification performance.

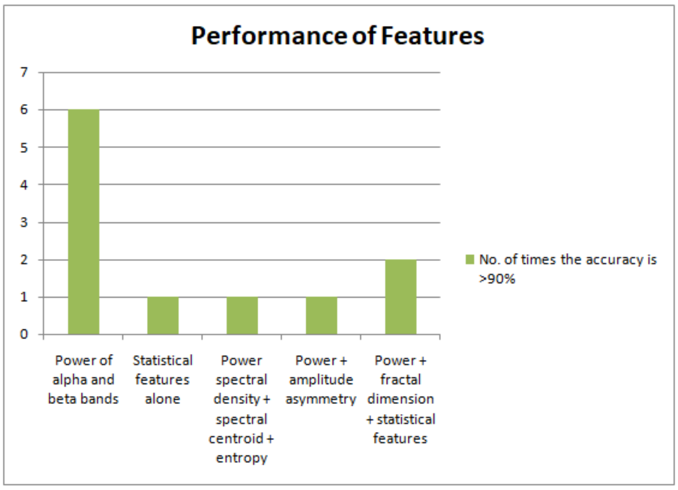


Figure 3.4: Performance of features in classification

The power of Alpha and Beta bands of EEG have been able to achieve an accuracy of 90% and above in 6 of the papers. These are frequency domain features, derived from Power Spectral Density (PSD) of the signal. The performance of rest of the features are represented in figure 3.4.

3.2.5 Preprocessing Techniques

A raw EEG signal captured using an EEG device contains a lot of noise and artefacts. These artefacts are eye blinking, head movement, muscle movement, twitching, etc. In order to get rid of these unwanted noise, filtering is used as a popular pre-processing technique. Filtering gets rid of the noises, which are of higher frequencies and allow the brain EEG signals of lower frequencies to pass through and stay. The method of decomposition of the EEG wave into its constituent frequency bands, normalization of data etc. are also considered as preprocessing technique. While most publicly available datasets are filtered and cleaned, preprocessing techniques will be of significance while working towards development of a stress feedback mechanism while researching with actual individuals as subjects and working with EEG devices.

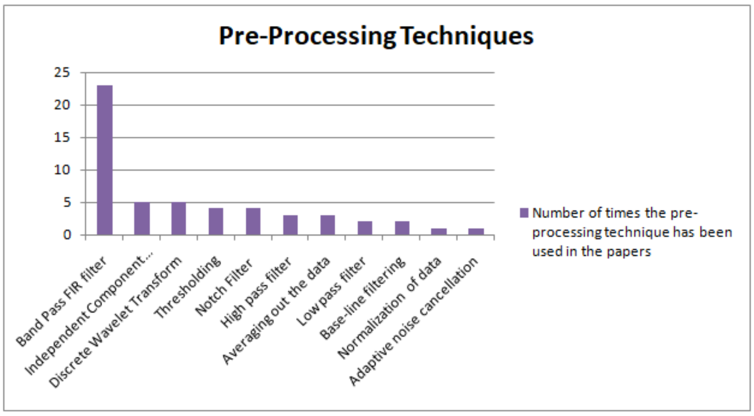


Figure 3.5: Choice of Preprocessing Techniques

From figure 3.5, we see that bandpass filter is an outright popular choice for filtering the EEG signals, being used 23 times. This filter can be customised to let a particular range of frequencies to pass (say 0.5 Hz - 45 Hz), while eliminating the noise having higher or lower frequencies.

In the next chapter, we discuss our work on the domain of EEG-based Emotion Recognition in detail. Emotion Recognition has been a fairly popular task that has often been addressed in cognitive neuroscience. We explain our methodologies in detail and mention our findings from the study that motivate us further as we go forward.

Chapter 4

An Approach to EEG-based Emotion Recognition using DEAP Dataset

Emotion recognition has been a popular application of machine learning and deep learning in recent years. We take up this task as a preliminary study towards working and understanding EEG signals for development of a mobile feedback mechanism. There are a number of reasons for taking up this task and its relevance to stress detection studies-

- The brain regions that are active during emotion and excitement are similar to the ones involved in stress detection studies. The frontal region of the brain is reported to be active during both emotional activities as well as stressful tasks [43, 44].
- The features that have performed well for emotion recognition studies have also performed well for stress detection tasks [27, 45].
- The two tasks are so much intertwined that the datasets that were originally developed for emotion analysis (DEAP [16], SEED [46]) were also explored and utilised in stress detection studies too [31, 47, 48].

We explored and tried to tackle this problem prior to working exclusively towards stress detection. Our goal was to develop a model that could be used for emotional state feedback, and help towards a stress detection model as well.

4.1 The model for emotion recognition

The circumplex model of emotion was conceptualized by James Russell and Lisa Feldmann Barett. This model suggests that emotions can be represented on a two-dimensional plane, with arousal and valence as the dimensions. Arousal represents the vertical axis and valence represents the horizontal axis, while the centre of the circle represents a neutral valence and a balanced level of arousal. Valence dimension represents how pleasant or unpleasant a feeling is. A happy person would fall at the positive end of the X-axis and an unhappy one on the negative end. Arousal dimension represents how exciting, arousing or stimulating an activity can be. An excited emotion would fall at the positive end of Y-axis, while a calming emotion would be represented on the negative end.

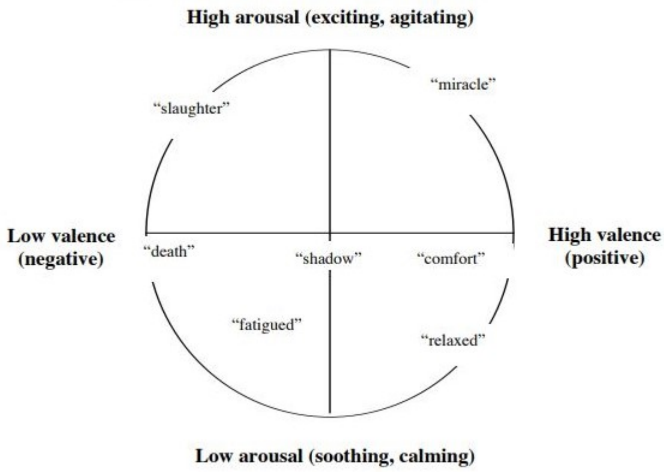


Figure 4.1: The Valence-Arousal Model for emotion recognition

4.2 Dataset

For training of the classification model, we have used the publicly available DEAP dataset [16]. The 'DEAP: A Database for Emotion Analysis using Physiological Signals' presents a multimodal dataset for the analysis of human emotion levels. The EEG and peripheral physiological signals of 32 volunteers who participated in the study were recorded as each of them were made to watch 40 one-minute long clips of music videos. Participants rated each video on their levels of arousal, valence, like/dislike, dominance and familiarity separately. The data was collected using a 32-channel EEG device. A participant in the study is shown in figure 4.2

Figure 4.2: A participant in the DEAP study

4.3 Feature selection

For creating the feature vector, we use bandpower- a frequency domain feature based on Power Spectral Density (PSD). PSD-based features have alternately performed well in emotion recognition and stress detection tasks. We divide the EEG data from each EEG channel into five fequency bands- Delta, Theta, Alpha, Beta, Gamma. We then find the average power over each of these five bands, also known as bandpower of the concerned frequency band, for each trial of every subject. We get a feature vector of dimension 1280x160.

4.4 Classifiers

For the classification of the data, we used the classic machine learning algorithms- K-Nearest Neighbours (KNN), Support Vector Machines (SVM), Decision Tree and Random Forest. Like the stress detection studies, SVM and KNN have also been able to perform significantly in emotion recognition studies. In addition to this, we also use a deep learning model, Long Short Term Memory (LSTM) which has been a popular choice of deep learning classifier in emotion recognition tasks in recent years.

Each of the 32 subjects had 40 trials (40 videos). So we developed two seperate models of classification- a subject-dependent model where the training and testing is done on the trials of the same subject, and an subject-independent model where the training and testing is done on all the trials irrespective of the subjects.

4.5 Results

We highlight the results obtained for the models in the table below:

Table 4.1: Testing accuracies for Subject-Dependent and Subject-Independent models

Model	Subject-Dependent		Subject-Independent	
	Valence	Arousal	Valence	Arousal
KNN	86.03	79.64	70.86	68.36
SVM	76.56	72.66	**72.19**	**71.25**
Decision Tree	71.10	67.97	58.13	55.63
Random Forest	81.25	81.19	61.95	61.25
LSTM	**94.69**	**93.13**	70.31	69.53

From the results, we have seen that the LSTM classifier outperforms the other classifiers by a significant margin- highlighting the reason why it has been a popular choice for classification in emotion recognition tasks in recent years.

4.6 Findings and Limitations

We achieved a significant accuracy while performing the classification on the DEAP dataset. We obtained a few findings that motivated the further studies in stress detection. We enumerate them below.

1. Frequency domain features, like in other EEG tasks, have performed well in terms of classification accuracy. The investigation with features can be narrowed down to the ones with frequency domain features only.

2. In presence of multiple trials for a single subject, a subject-dependent model will give a better accuracy than a subject-independent model. This is because the EEG signals vary greatly from person to person, while they are extremely alike for a single person even across multiple trials.

3. The LSTM can prove to be a powerful tool for subject-dependenct studies. The solution to the problem of long-term dependencies can be provided by the LSTM model in a way that no other RNN model does. This is beneficial for subject dependent studies where significant information can be retained by the LSTM model during training over multiple trials of the same subject.

4. The LSTM model fails to retain the same performance in subject-independent study. The property of LSTM to retain information over longer training duration hinders the performance in a subject-independent strategy where the trials vary greatly and have no significant information to be derived over different trials of different subjects.

5. Out of all the classifiers, the performance of SVM remains relatively same over both subject-dependent and subject-independent strategies. Thus, SVM can be an effective tool for classification over new, unknown or subject independent data. This is consistent with our findings from the literature review in the earlier chapter.

6. Wichakam and Vateekul [45] in their emotion recognition study reports that the accuracy of a classification model can be improved using fewer EEG channels. This motivates us to look for narrowing down our future study to fewer and relevant EEG channels only. Feature selection methods are capable of providing us with the relevant and important features and EEG channels for our study.

4.7 Publications

Parts of this work have now been published as two research papers:

1. Debarshi Nath, Anubhav, Mrigank Singh, Divyashikha Sethia, Diksha Kalra, S. Indu, "*A Comparative Study of Subject-Dependent and Subject-Independent Strategies for EEG-Based Emotion Recognition using LSTM Network*", in ACM Proceedings of 4th International Conference on Compute and Data Analysis (ICCDA 2020), March 9-12, 2020, pp. 142-147

2. Anubhav, Debarshi Nath, Mrigank Singh, Divyashikha Sethia, Diksha Kalra, S. Indu, "*An Efficient Approach to EEG-Based Emotion Recognition using LSTM Network*", in IEEE Proceedings of 16th IEEE Colloquium on Signal Processing and its Applications (CSPA 2020), February 28-29, 2020, pp. 88-92

Chapter 5

Mental Workload Detection using EEG Signals from "EEG During Mental Arithmetic Tasks" Dataset

Our study and involvement in the area of EEG-based emotion recognition drive us and motivates us towards a related yet unlike application of EEG and machine learning- short-term stress detection. Long term mental stress can lead to serious mental disorder, and repeated small-term stress and anxiety is a roadway towards these long-term problems. With our experience from the previous task of emotion recognition, we identify a few immediate objectives for our study-

1. To find the effectiveness of a small dataset of short duration in making assessment of mental workload.
2. Identify the specific frequency domain features that can provide good classification for the task
3. Compare the performance of the classifiers identified from the literature review for the particular task.
4. Identify the brain regions that are more active during the task and improve classification.
5. Contribute towards the long-term goal of establishing a concrete model for stress detection in people using feedback from physiological sensors.

We outline our study in the following sections.

5.1 Dataset

In this work, we analyse the publicly available EEG During Mental Arithmetic Tasks [17] dataset.

All the participants in the study were reported fit with no history of cognitive or mental impairment or any verbal or non-verbal learning disabilities. The EEG data were collected while the participants performed the arithmetic task in which they calculate serial subtraction of two numbers. For each trial, the 4 digit minuend and the 2 digit subtrahend was communicated orally. The complete procedure for the data recording was over 10 minutes, including both rest and mental activity phases.

After making the participants adapt with the testing conditions, there is 3 min. of rest phase, ensuring that the participant is calm before the task. The following test involves a mental arithmetic task phase of 4 minutes duration. This dataset reports EEG and ECG recordings for the complete rest phase (3 min.) and the first minute of mental activity phase. During long mental tasks, participants are more susceptible to fatigue and emotional workload, that's why considering data for the first minute is justified.

This dataset reports the EEG data of 36 participants recorded using the Neurocom EEG 23-channel system (Ukraine, XAI-MEDICA). This device has a sampling frequency of 500 Hz, and the placement of electrodes is according to the 10-20 International system covering the Anterior Frontal, Frontal, Central, Temporal, Parietal, and Occipital regions of the brain. A High pass filter of 0.5Hz cutoff frequency and a low pass filter with cutoff frequency 45Hz along with a 50Hz notch filter were employed for artefact removal. Also, ICA was applied to remove the artefacts resulting from eye blinks, muscle movements and cardiac overlapping of cardiac pulsation.

5.2 Features

For creating the feature vector for the classifiers, we used the following four features (features discussed in detail in Chapter 2).

1. Spectral Entropy
2. Relative Power
3. Alpha-Beta Ratio
4. Theta-Alpha Ratio

5.3 Feature Selection

Feature selection is the analytical process of selecting a combination of significant features that can improve the performance in classification. Selecting only the significant features has numerous benefits for developing an efficient model, such as:

- Dimensionality Reduction
- Faster computation
- Simpler predictive models
- Generalisation

Our another objective of applying feature selection techniques is to find out the brain regions, and thus the EEG channels, that are most involved in mental workload tasks. We also want to verify the results and findings of the earlier studies in this avenue. We select three feature selection techniques for our study, described below.

1. **ANOVA**: Analysis of Variance (ANOVA) is a statistical model that contrast the differences of the group means in the data. ANOVA is a filter method of feature selection. It forms the following hypotheses and tests them for the data sample:

   - *Null Hypothesis*: All groups have equal variance.
   - *Alternative Hypothesis*: All groups have unequal variance.

   Results are considered statistically significant if p-value (probability value) is less than the threshold; thus, rejecting the null hypothesis indicating that there exists some variance between the groups. The results of the F-test rank the features according to their importance. Therefore, we can select a subset of features accordingly, making the analysis of differences in the activation of distinct regions of the brain easier.

2. **RFECV**: Recursive Feature Elimination using Cross-Validation (RFECV) is a wrapper method for feature ranking and selection. It is a slight modification of Recursive Feature Elimination (RFE) introduced by Guyon et al. [49] in 2002. Authors describe an increment of 12% to the testing accuracy for the significant features obtained using RFE.

   In RFE, the estimator first trains on the complete set of features then, recursively select the best features according to feature importance, i.e. weight coefficient for the classifier until the criterion for minimum features is met. While RFECV, implements cross-validation for training the estimator, this ensures the selection of robust features. This method gives the set of the best features for the classifier, but it can be computationally expensive.

3. **Boruta**: Boruta, introduced by Kursa and Rudnicki [50], is an embedded method of feature selection. This algorithm integrates the feature selection along with the training process of the Random Forest classifier. Thus, repetitively rejecting features based on their statistical significance. Boruta was first implemented in the R language by the author.

5.4 Methodology

We implement SVM, KNN, Decision Tree, and Naive Bayes classifiers using scikit-learn module [51] of Python 3 on the Google Colab platform. EEG during Mental Arithmetic Tasks dataset reports data in European Data Format (EDF) file format. We use the open-source module PyEDFlib [52] in Python 3 to handle EDF files. We extract the EEG features and formulate a classification problem to identify rest v/s mental activity phase for all the participants collectively. We label the features from rest phase as (0) and mental activity phase as (1). The flowchart for the modelling process is illustrated in figure 5.1.

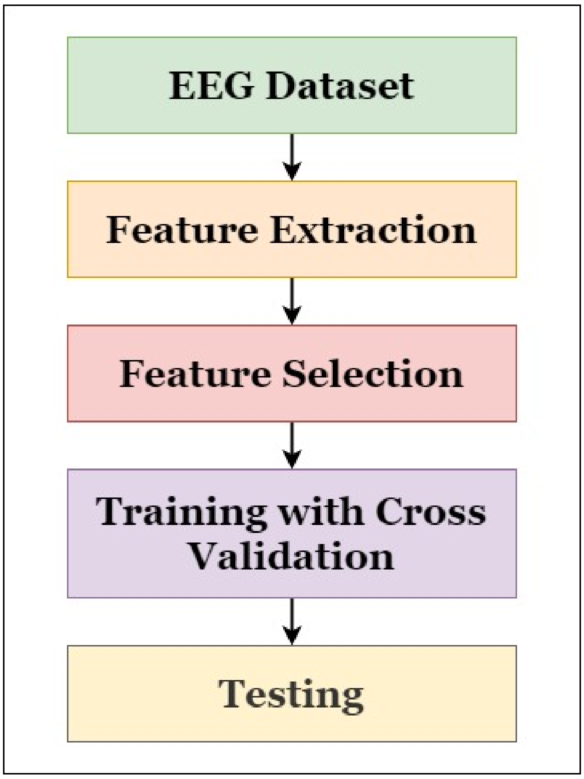


Figure 5.1: Flowchart for classification of rest v/s mental activity phase

To ensure unbiased training of models, we use the Stratified K-fold cross-validation method for training and testing the classifiers. Using the Stratified folds ensures that all the subsets are balanced, having equal instances of both classes. We use the value of $k = 6$ in our analysis. To search the best configuration for the classifiers, we implement the GridSearchCV from scikit-learn. This method constructs a grid of hyper-parameters and selects the best combination of hyper-parameters using the cross-validation method. We enlist the explored hyper-parameters in Table 5.1.

Table 5.1: Range of parameters for different classifiers

Classifier	**Parameters**
SVM	C:$[0.01, 0.1, 1, 10, 100]$, kernel:['linear', 'rbf']
Decision Tree	criterion:['entropy', 'gini'], max_depth:$[5, 10, 15, 20, None]$
KNN	n_neighbors: $[3, 5, 7, 9, 11]$, distance:['euclidean', 'manhattan', 'chebyshev']

5.5 Results and Analysis

The classification results for the rest v/s mental activity phases of participants are examined using the classifiers namely Naive Bayes, Support Vector Machine (SVM), Decision Tree, and K-Nearest Neighbor (KNN). We compare the Accuracy and F1 score of these classifiers for the features used for training. We also contrast the results for the subset

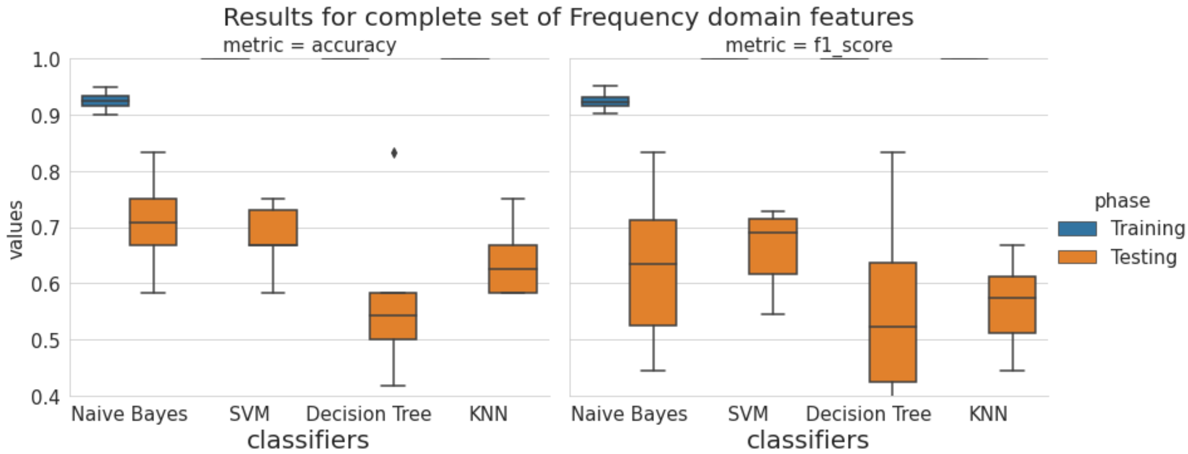
Table 5.2: Classification result for different classifiers trained on Frequency domain features

Classifier	Feature set	Training		Testing	
		Accuracy(%)	F1 score(%)	Accuracy(%)	F1 score(%)
Naive Bayes	Complete	92.5 ± 1.6	92.44 ± 1.53	70.83 ± 7.98	62.86 ± 13.19
	ANOVA	66.94 ± 3.10	65.52 ± 2.80	62.5 ± 14.23	59.63 ± 15.74
	RFECV	$\mathbf{89.44 \pm 2.66}$	$\mathbf{89.14 \pm 2.72}$	$\mathbf{77.78 \pm 8.89}$	$\mathbf{78.67 \pm 8.49}$
	Boruta	68.06 ± 1.79	63.87 ± 2.46	68.05 ± 11.2	61.65 ± 13.84
SVM	Complete	100 ± 0	100 ± 0	68.06 ± 5.73	66.13 ± 6.73
	ANOVA	69.72 ± 3.65	70.48 ± 3.30	66.67 ± 12.73	67.4 ± 10.39
	RFECV	$\mathbf{99.17 \pm 1.27}$	$\mathbf{99.14 \pm 1.31}$	$\mathbf{93.05 \pm 8.89}$	$\mathbf{93.36 \pm 8.26}$
	Boruta	75.28 ± 4.02	75.16 ± 6.52	69.44 ± 15.71	67.65 ± 21.10
Decision Tree	Complete	100 ± 0	100 ± 0	56.94 ± 13.10	52.79 ± 19.3
	ANOVA	99.17 ± 1.27	99.18 ± 1.24	58.33 ± 18.0	60.2 ± 18.09
	RFECV	99.17 ± 1.86	99.12 ± 1.96	65.28 ± 10.11	64.09 ± 9.68
	Boruta	$\mathbf{88.33 \pm 4.19}$	$\mathbf{89.1 \pm 3.64}$	$\mathbf{69.44 \pm 14.16}$	$\mathbf{71.43 \pm 11.76}$
KNN	Complete	100 ± 0	100 ± 0	63.89 ± 6.21	56.2 ± 7.44
	ANOVA	70.29 ± 13.92	67.88 ± 14.75	62.5 ± 13.39	58.00 ± 13.15
	RFECV	100 ± 0	100 ± 0	61.11 ± 14.16	46.06 ± 17.57
	Boruta	$\mathbf{80.28 \pm 9}$	$\mathbf{81.66 \pm 8.36}$	$\mathbf{70.83 \pm 14.23}$	$\mathbf{73.97 \pm 10.94}$

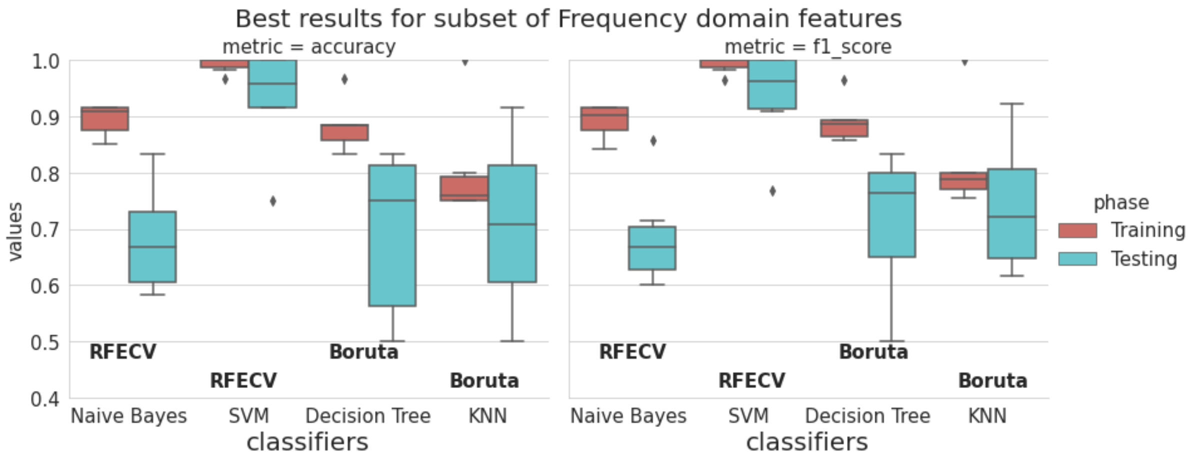
of features selected by different feature selection techniques. We show the classification results for all the feature selection methods in

Figure 5.2 depicts the results using the Frequency domain features. Naive Bayes performs best with 70.83% testing Accuracy while the other classifiers suffer heavily from over-fitting for the complete set of features, in Figure 5.2 (a). We note a significant improvement in the results using the feature selection methods, Figure 5.2 (b) depicts the highest gain of 25% in testing Accuracy and F1 score for the SVM classifier as it obtains highest testing Accuracy and F1 score of 93.06% and 93.35%, this verifies the need for using feature selection.

Now we wish to look at which features had been selected most frequently by the feature selection methods. We visualize the frequency of the selection of these features through figure 5.3.



(a)



(b)

Figure 5.2: Box plots of Accuracy and F1 scores obtained using the Frequency-domain features for rest v/s mental activity phase. Results from Stratified K-fold training of the classifiers, namely Naive Bayes, SVM, Decision Tree, and KNN, are used in the plot. (a) outlines the results for the complete set of Frequency-domain features, we note the best testing results for the Naive Bayes classifier. (b) illustrates the best of the results observed for the classifiers trained on the subsets of features. The optimum selection method is also indicated such as SVM acquires best results for the features selected by RFECV.

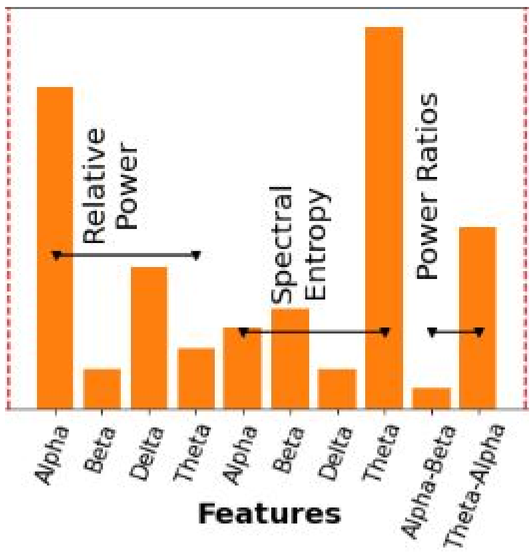


Figure 5.3: Frequency of the features selected by the feature selection algorithms

From the figure 5.3, we can that the highest selected features from the group of frequency-domain features are the alpha relative power and the spectral entropy of the theta band. Theta-Alpha Ratio which had been used as a metric for recognizing mental workload in earlier studies is also selected with considerable frquency.

The frequency of selection of each of the channels of EEG can give an indication of which brain regions are to be targeted for a stress detection study. We look at the figure 5.4 for the frequency of selection for these channels.

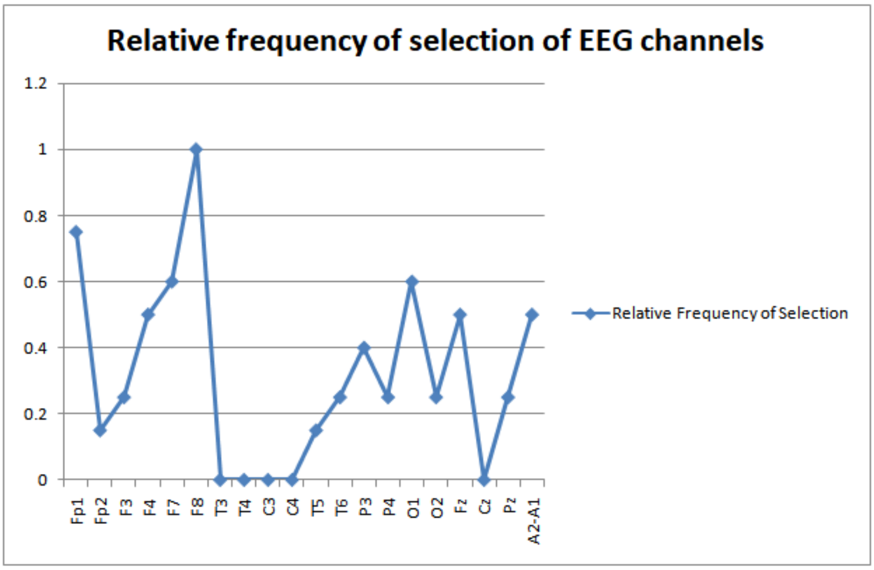


Figure 5.4: Frequency of the EEG channels selected by the feature selection algorithms

From figure 5.4, we see that the frontal channels have the higher selection rate, which is consistent with earlier studies. The occipital regions too have been selected a considerable number of times.

5.6 Findings and Limitations

The highlight a few of the findings that we have found from analysis of the EEG During Mental Arithmetic Task dataset:

1. Naive Bayes performed the best for complete set of features.
2. SVM classifier, along with RFECV feature selection improved the accuracy of the classifier by almost 23%, upto 93.05% accuracy.
3. The theta and alpha bands are more frequently selected by feature selection algorithms.
4. The frontal EEG channels of the brain have been most frequently selected by the feature selection algorithms.
5. Some of the classifiers have shown high overfitting. There is a need for a larger EEG database to translate the findings of EEG studies into real world.
6. To have a subject-specific personalised stress detection system, it is necessary to have multiple trials of the same subject. This is also shown by our study on EEG-based Emotion Recognition in Chapter 4.

5.7 Publications

This study has been submitted in form of a research paper, *Cognitive Workload Detection in a Mental Arithmetic Task: An Investigative Study of EEG-Features and Feature Selection Methods* in the journal 'Cognitive Systems Research'. The research paper is currently under review.

Chapter 6

Cognitive-Mental Assessment using Physiological Signals: A Stress Detection Study on Real Participants

In the earlier chapter, we tested the classifiers on a dataset of EEG signals recorded during a mental arithmetic task. We achieved a significantly good classification accuracy, outperforming state-of-the-art classifiers in the process on that dataset. We found out the significance of the frontal EEG channels in the assessment of mental workload generated through a mental arithmetic task.

In the task described in this chapter, we perform a similar experiment- along with making use of the findings of the task described in the previous chapter. We make use of EEG devices to obtain signals from subjects taking part in our experiments, and clean and preprocess the data, and use them for a classification task. To reiterate the long term goal of the project- we want to have a mechanism to detect stress with the feedback of physiological signals from portable mobile devices that may be validated by EEG signals. We title the project 'Cognitive-Mental Assessment using Physiological Signals' or 'Cog-MAPS'.

6.1 Equipments Required

For this experiment, we require an EEG device. At Samsung Digital Academy Research lab, we have 'Emotiv EPOC+', a 14-channel portable EEG headset. The device is easy to handle, convenient, and provides a software for its Brain-Computer Interface for post recording analyses.

Figure 6.1: The Emotiv EPOC+ EEG Device (two units)

The other device that we have at our disposal is the E4 Empatica wristband. Otherwise looking just like a wristwatch, the medical grade E4 Empatica wristband provides us with 4 physiological sensors-

1. **Photoplethysmograph (PPG) Sensor**: Measures volumetric changes in blood in peripheral circulation.
2. **Galvanic Skin Response (GSR) Sensor**: Measures the Electrodermal Activity (EDA) of the skin, a measure of the perspiration rate.
3. **Infrared Thermopile**: Measures the peripheral skin temperature.
4. **Accelerometer**: Captures the motion based activity.

Figure 6.2: The E4 Empatica Wristband

Apart from the two devices, the participants we also used two questionnaires to be filled by the participants in the study. The purpose of the questionnaires was to validate and assess the stress states of the individuals before and after the test. These questionnaires are to be made available for further studies, should anyone wishes to investigate further using the dataset.

The first of the questionnaires is the Perceived Stress Scale (PSS) [3]. The motive of this questionnaire is to assess the perceived stress (stress due to long-term social and personal issues like failing career, death of a relative, etc.) of the individual at that time. The questionnaire is shown in Figure 6.3.

PERCEIVED STRESS SCALE

The questions in this scale ask you about your feelings and thoughts during the last month. **In each case, you will be asked to indicate by circling *how often* you felt or thought a certain way.**

Name ______________________ Date ____________

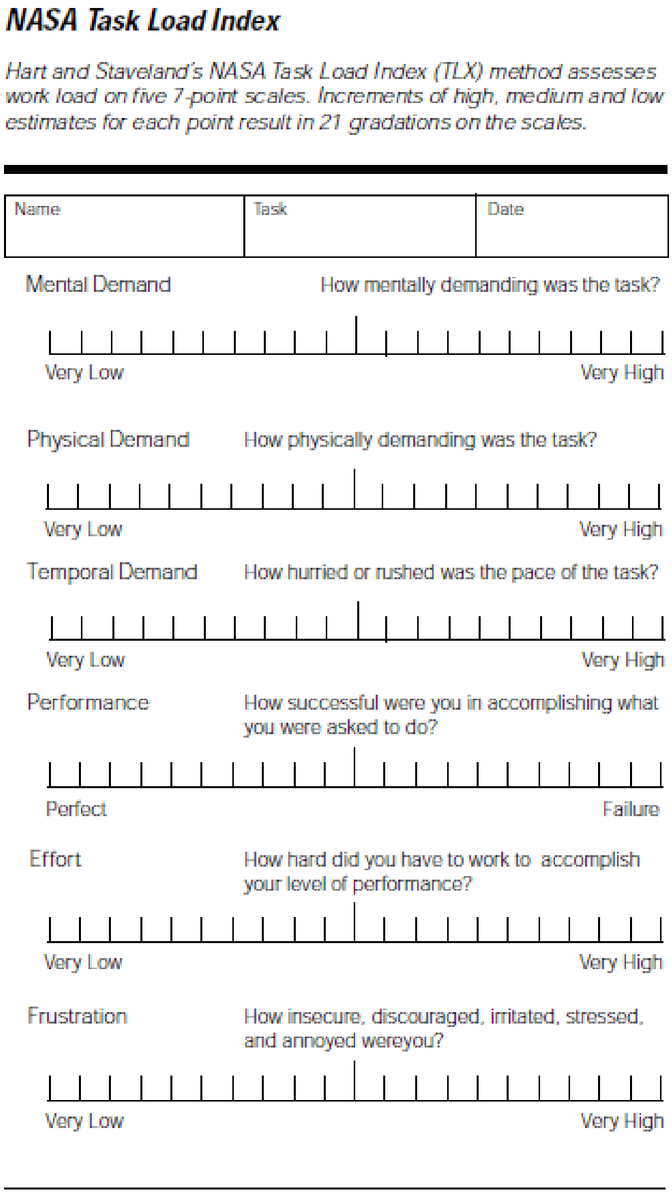
Age _______ Gender (*Circle*): **M F** Other ______________________

0 = Never 1 = Almost Never 2 = Sometimes 3 = Fairly Often 4 = Very Often

1. In the last month, how often have you been upset because of something that happened unexpectedly?	0	1	2	3	4
2. In the last month, how often have you felt that you were unable to control the important things in your life?	0	1	2	3	4
3. In the last month, how often have you felt nervous and "stressed"?	0	1	2	3	4
4. In the last month, how often have you felt confident about your ability to handle your personal problems?	0	1	2	3	4
5. In the last month, how often have you felt that things were going your way?	0	1	2	3	4
6. In the last month, how often have you found that you could not cope with all the things that you had to do?	0	1	2	3	4
7. In the last month, how often have you been able to control irritations in your life?	0	1	2	3	4
8. In the last month, how often have you felt that you were on top of things?	0	1	2	3	4
9. In the last month, how often have you been angered because of things that were outside of your control?	0	1	2	3	4
10. In the last month, how often have you felt difficulties were piling up so high that you could not overcome them?	0	1	2	3	4

Figure 6.3: The Perceived Stress Scale (PSS) Questionnaire [3]

The second questionnaire is The NASA Task Load Index (NASA-TLX) questionnaire [4]. This questionnaire has been popularly used in order to measure short-term stress as a feedback from participants. The questionnaire is shown in Figure 6.4.



NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task	Date

Mental Demand — How mentally demanding was the task?

Very Low — Very High

Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?

Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?

Very Low — Very High

Figure 6.4: The NASA Task Load Index (NASA-TLX) Questionnaire [4]

We also used a webcam in order to capture the movements of the subject in order to later validate motion artefacts from these videos.

6.2 Experimental Setup

To acquire participants for the study, we put up posters and flyers around the DTU college campus. A notification was put up on the DTU official website "www.dtu.ac.in" and a sepearate webpage was created for the program on the SDA Research Lab Website.

For the study, a cardboard cubicle was created of dimensions 3ft x 3ft (approx.). The idea was to have minimum interference and noise at the test area. The cubicle was covered with white wallpapers to have a well-lit area.

The Montreal Imaging Stress Task (MIST) is a fairly recent and popular mental arithmetic test which has been extensively used in stress studies [53].To conduct the test, we make use of Inquisit Lab which is a powerful tool that allows designing interactive psychological tests. A ready-made, editable version of the Montreal Imaging Stress Task (MIST) test is provided by Millisecond [54]. It allows a script to be run on a computer, and a participant can interactively participate in the test. A snap of the test with description within is provided in figure 6.5.

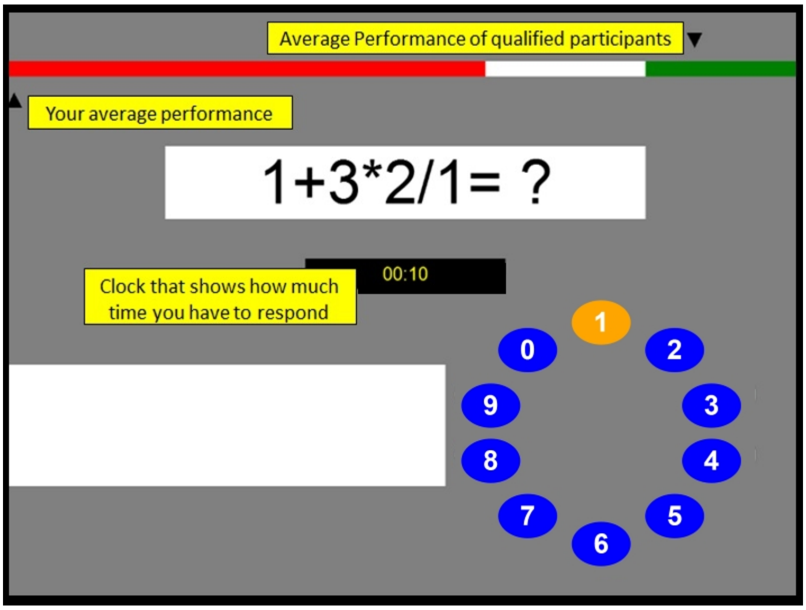


Figure 6.5: A snapshot of the MIST test interface

6.3 Methodology

6.3.1 Test Procedure

The participant was scheduled for the test at a mutually agreed date and time. The participant was instructed not to consume coffee or caffeine at least within four hours prior to the test as caffeine is known to highly interfere with EEG [55]. Participants were asked not to apply hair oil to ensure proper contact with EEG electrodes. Till the time of writing this thesis, we could collect the data of 5 individuals. The MIST test duration was approximately 40 mins, while the entire test procedure was completed in 1 hour for each participant.

The procedure followed on the day of the test is listed in the following steps-

- Participants arrive, and are made to sit and listen to the test procedure.
- They are made to fill a consent form and PSS Stress Scale Form.
- The devices are put on them.
- The participant moves to the test area.
- The test consists of six major phases, chronologically listed below:
  1. Training: The participant gets familiar with the test interface.

2. Rest: 5 mins of inactivity where the participant is asked to relax.
3. Control: The participant takes the untimed MIST task.
4. Timed Test: The participant takes the timed MIST task with feedback.
5. Control: The participant takes the untimed MIST task.
6. Rest: 5 mins of inactivity where the participant is asked to relax

- The devices are removed.
- The participant is asked to fill the NASA-TLX Form
- End

6.3.2 Data Preprocessing

The data was then exported from the devices using their respective softwares, EmotivPRO for the Emotiv EPOC+ EEG device and E4 Manager for the E4 Empatica wristband. Working solely with the EEG data for the remainder of the project, the EEG signals were read in Python 3.0, and cleaned using a bandpass filter of 0.5Hz - 45 Hz, removing the artefacts outside this frequency range.

From these EEG signals, we use only the frontal EEG channels, i.e., F7, F3, F4, F8 for our task of stress detection. We use frontal channels from our findings from the work described in previous chapter where we found the effectiveness of the frontal region of brain in detecting mental workload.

Each EEG signal was then cut into three shorter signals. These three portions of the signal are from three different phases of the MIST task:

1. Rest Phase: The first rest phase where the participant is asked to rest while keeping eyes open. (Signal duration: 3 minutes)
2. Control Phase: The participant performs the untimed MIST Task. (Signal duration: 2 minutes)
3. Test Phase: The participant performs the timed MIST Task. (Signal duration: 2 minutes)

From these portions of the signal, we formulate a classification model or stress detection task- identifying which phase of the task the signal is from. Alternately, we identify and detect the amount of mental workload and stress in the individual.

6.3.3 Feature Extraction

From the signals we generate the same four frequency domain features that we investigated in the previous chapter-

1. Spectral Entropy
2. Relative Power
3. Theta-Alpha Ratio

4. Alpha-Beta Ratio

Prior to finding out these features, we separate each of the EEG signal into the five frequency bands- Delta, Theta, Alpha, Beta and Gamma.

The data that we could collect from the experiment is limited. During the data collection phase, we could manage to get the data from 5 subjects. From each of these 5 subjects, we get the signals from the three phases mentioned above.

6.3.4 Problem Formulation

We use these features to formulate two different classification problems-

1. 2-class stress detection: A two class classification problems to identify the rest phase and the test phase from the signals.

2. 3-class stress detection: A three class classification to identify the rest, control and the test phases from the signals.

For the 2-class classification, we have two phases of signals for five subjects. The feature vector, hence, is of dimensions 48x10. For the 3-class classification, the feature vector is of size 48x15. A visual representation of the methodology is given in figure 6.6.

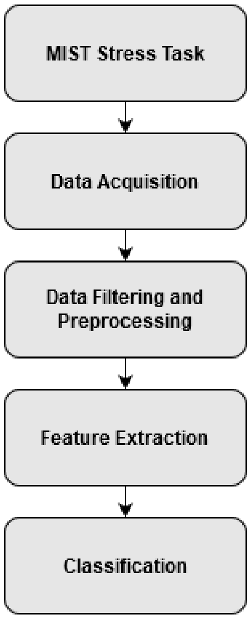


Figure 6.6: Flowchart for the classification model

All the experiments were done on the Google Colab Platform with implementation in Python 3.0. PyEDFLib was used to read the signals from EDF format, the mne package was used to design the bandpass filter, and scikit-learn to implement the machine learning models. We use Grid Search cross validation for hyperparameter tuning. Principal Component Analysis was performed for dimensionality reduction. Cross validation is done using StratifiedKFold Cross Validation with 5 folds.

6.4 Results

The best results of the classification are summarized in table 6.1.

Table 6.1: Classification accuracies for the data obtained from the subjects

Classifier	2-class classification	3-class classification
SVM	90 ±20	53.33 ±26.67

The results obtained for the 2-class model is 90%. However there is a lot of deviation, signifying bias and a need for more data for training. The performance for 3-class classification is expectedly lower, with an accuracy of 53.33%. The results are comparable to the performance that we observed on the "EEG During Mental Arithmetic Tasks" dataset in Chapter 6.

6.5 Findings and Limitations

The findings from this particular study can be highlighted in the following points-

1. The model can be effectively used for stress detection. However, there is a need for more data.
2. The remaining physiological signals from E4 Empatica Wristband can be used in conjuction with EEG for a more effective classification model.
3. There is a need for more EEG data for building an effective stress detection model. Panicker et al. [56] suggests that at least 100 subjects are necessary in an EEG study to translate the results into real world.
4. The questionnaire measures can be integrated in the EEG study and be investigated further. Al-Shargie et al. [43] found that for multi-level stress assessment, the NASA Task Load Index (NASA-TLX) scale had no significant correlation with EEG stress levels. These reports can be investigated further.
5. The body and eye movements of the subjects captured using the webcam can be further explored for assessment of stress [12].

Chapter 7

Conclusion and Future Work

The works described in this thesis concerns with the goal of creating a robust and effective model for stress detection, using feedback from mobile sensors. Working towards it, we first tackle a popular and related problem in EEG applications- Emotion Recognition. After obtaining a good classification performance and numerous findings and revelations, we work towards our chief goal of stress detection. We work with a publicly available dataset that contains EEG signals recorded during a mental arithmetic task. Equipped with the findings of this study, we finally design a stress detection models for the EEG signals recorded from actual participants of a study in Samsung Digital Academy Research Lab, DTU. The chief contribution of this work is it proposes a model that can be used to acquire physiological data from subjects, analyse them and detect stress efficiently in that individual.

Despite considerable work in this area, there are still plenty of avenues to venture into to have results that are ultimately conclusive and definitive. Investigating the poor performance of some classifiers, we observe a need for a more extensive database. Thus, we aim to create a comprehensive database for stress study with more number of subjects and different customised cognitive tasks and protocols. This personalised data collection methodology will give more flexibility to EEG data and open new domains for analysis. One of such databases that can be worked towards is the one that was created at Samsung Digital Academy Research Lab, DTU.

The features relevant to the task of stress detection and classification are often task-dependent. So et al. [57] found that different cognitive activities result in a change in different EEG frequencies. The studies borrowing datasets from other studies are limited in terms of cognitive tasks that EEG analyses can be considered. Datasets are fixed, and we can only do the investigations on the limited amount of data that is available in the datasets. Panicker and Gayathri [56] reports in their survey that there is a need to create a database with preferably more than 100 subjects to have the findings of a study translate into the real world. Creating own database for stress studies with a sufficient number of subjects and different customised cognitive tasks and protocols for EEG data collection can give more flexibility and increase the number of analyses that can be done on a customised data collection methodology. Multiple cognitive tasks can be pipelined in the same protocol and used for stress detection studies as done by the works of Jun et al. [58] and Patel et al. [27].Other works like Thejaswini et al. [48] joins two datasets and use them for emotional state classification. This method of combining similar datasets can be useful for modelling with limited resources. Further, multiple trials for each subject open another avenue to explore into for the development of subject-dependent and subject-independent models for stress detection [59]. In case of availability of numerous EEG recordings for a single subject, the ability of LSTM Networks to efficiently capture long-term dependencies can be of significance in the development of a subject-specific stress detection model [60].

Real-time stress detection has a particular benefit in real-world scenarios as they are capable of reporting unsavoury stimuli in real-time. The work in our study can be extended to real-time situations, like many other real-time studies [61–63]. The crux of a real-time system is the ability to report its result immediately and quickly. Wang et al. [63] in their study demonstrates reducing the time complexity of the feature extraction process coupled with selecting a smaller subset of the EEG channels is capable of producing considerably robust and faster classification results. This is what was done in our study as well. We obtained a decent classification accuracy by using only four channels of the 14-channel Emotiv Epoc+.

Finally, when working toward stress recognition and classification, we must keep in mind the ultimate goal of these studies, which is stress reduction. While a considerable number of works has been done in the area of stress detection, the number of studies addressing stress reductions are fewer. There is a scope of developing methodologies for reduction of stress in conjunction with stress induction, like the ones shown by the studies of Patil et al. [64] and Kalas et al. [65].

Publications

The publications from the work in this thesis are:

1. Debarshi Nath, Anubhav, Mrigank Singh, Divyashikha Sethia, Diksha Kalra, S. Indu, "*A Comparative Study of Subject-Dependent and Subject-Independent Strategies for EEG-Based Emotion Recognition using LSTM Network*", in ACM Proceedings of 4th International Conference on Compute and Data Analysis (ICCDA 2020), March 9-12, 2020, pp. 142-147

2. Anubhav, Debarshi Nath, Mrigank Singh, Divyashikha Sethia, Diksha Kalra, S. Indu, "*An Efficient Approach to EEG-Based Emotion Recognition using LSTM Network*", in IEEE Proceedings of 16th IEEE Colloquium on Signal Processing and its Applications (CSPA 2020), February 28-29, 2020, pp. 88-92

3. Debarshi Nath, Anubhav, Diksha Kalra, Mrigank Singh, Divyashikha Sethia, S. Indu, "*Cognitive Workload Detection in a Mental Arithmetic Task: An Investigative Study of EEG-Features and Feature Selection Methods*", Cognitive Systems Research (under review)

Part of the work has also been presented at:

1. Poster Symposium in the "*International Workshop on Science of Intelligence*" at Indian Institute of Technology Jodhpur during January 18-19, 2020.

Bibliography

[1] A. Craik, Y. He, and J. L. Contreras-Vidal, "Deep learning for electroencephalogram (EEG) classification tasks: a review," *Journal of Neural Engineering*, vol. 16, no. 3, p. 031001, Sep 2019.

[2] R. Vallat. Bandpower of an EEG Signal. [Online]. Available: https://raphaelvallat.com/bandpower.html(accessedNov.13,2019)

[3] S. Cohen, T. Kamarck, and R. Mermelstein, "A Global Measure of Perceived Stress," *Journal of Health and Social Behavior*, vol. 24, no. 4, p. 385, 1983.

[4] S. G. Hart and L. E. Staveland, "Development of nasa-tlx (task load index): Results of empirical and theoretical research," in *Human Mental Workload*, ser. Advances in Psychology, P. A. Hancock and N. Meshkati, Eds. North-Holland, 1988, vol. 52, pp. 139 – 183. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0166411508623869

[5] L. Dubreuil, "How can we apply AI, Machine Learning or Deep Learning to EEG?" Mar 2018. [Online]. Available: https://www.neuroelectrics.com/blog/from-ai-to-deep-learning-applied-to-eeg

[6] J. Alvey, "Using deep learning to "read your thoughts"-with keras and an EEG sensor," Apr 2019. [Online]. Available: https://medium.com/@justlv/using-ai-to-read-your-thoughts-with-keras-and-an-eeg-sensor-167ace32e84a

[7] T. D. Fulvio, "Using machine learning to categorise EEG signals from the brain to words," Apr 2019. [Online]. Available: https://towardsdatascience.com/using-machine-learning-to-categorise-eeg-signals-from-the-brain-to-words-728aba93b2b3

[8] S. Tan and A. Yip, "Hans Selye (1907–1982): Founder of the stress theory," *Singapore Medical Journal*, vol. 59, no. 4, p. 170–171, 2018.

[9] M. D. Breed and J. Moore, "Cognition," *Animal Behavior*, p. 151–182, 2012.

[10] M. J. Horowitz and S. S. Becker, "Cognitive Response to Stressful Stimuli," *Archives of General Psychiatry*, vol. 25, no. 5, pp. 419–428, 11 1971.

[11] "Stress and cognitive workload," in *Tactical display for soldiers: human factors considerations.* National Academy Press, 2007.

[12] G. Giannakakis, D. Grigoriadis, K. Giannakaki, O. Simantiraki, A. Roniotis, and M. Tsiknakis, "Review on psychological stress detection using biosignals," *IEEE Transactions on Affective Computing*, pp. 1–1, 2019.

[13] A. de Santos Sierra, C. Sánchez Ávila, J. Casanova, and G. Bailador, *Real-Time Stress Detection by Means of Physiological Signals*, 07 2011, pp. 23–44.

[14] "What are brainwaves?" [Online]. Available: https://brainworksneurotherapy.com/what-are-brainwaves

[15] Y. Roy, H. Banville, I. Albuquerque, A. Gramfort, T. H. Falk, and J. Faubert, "Deep learning-based electroencephalography analysis: a systematic review," *Journal of Neural Engineering*, vol. 16, no. 5, p. 051001, 2019.

[16] S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "DEAP: A Database for Emotion Analysis; Using Physiological Signals," *IEEE transactions on affective computing*, vol. 3, no. 1, pp. 18–31, 2011.

[17] I. Zyma, S. Tukaev, I. Seleznov, K. Kiyono, A. Popov, M. Chernykh, and O. Shpenkov, "Electroencephalograms during Mental Arithmetic Task Performance," *Data*, vol. 4, 2019.

[18] P. Welch, "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," *IEEE Transactions on Audio and Electroacoustics*, vol. 15, no. 2, pp. 70–73, 1967.

[19] Y. Tian, H. Zhang, W. Xu, H. Zhang, L. Yang, S. Zheng, and Y. Shi, "Spectral Entropy Can Predict Changes of Working Memory Performance Reduced by Short-Time Training in the Delayed-Match-to-Sample Task," *Frontiers in Human Neuroscience*, vol. 11, p. 437, 2017. [Online]. Available: https://www.frontiersin.org/article/10.3389/fnhum.2017.00437

[20] C. Vuppalapati, N. Raghu, P. Veluru, S. Khursheed *et al.*, "A system to detect mental stress using machine learning and mobile development," in *Proc. of International Conference on Machine Learning and Cybernetics (ICMLC)*, vol. 1. IEEE, 2018, pp. 161–166.

[21] A. R. Subhani, W. Mumtaz, M. N. B. M. Saad, N. Kamel, and A. S. Malik, "Machine learning framework for the detection of mental stress at multiple levels," *IEEE Access*, vol. 5, pp. 13 545–13 556, 2017.

[22] A. R. Subhani, W. Mumtaz, N. Kamil, N. M. Saad, N. Nandagopal, and A. S. Malik, "MRMR based feature selection for the classification of stress using EEG," in *Proc. of Eleventh International Conference on Sensing Technology (ICST)*. IEEE, 2017, pp. 1–4.

[23] A. Niemiec and B. Lithgow, "Alpha-band characteristics in EEG spectrum indicate reliability of frontal brain asymmetry measures in diagnosis of depression," in *Proc. of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*. IEEE, Institute of Electrical and Electronics Engineers, 2005, pp. 1 – 4.

[24] G. Sammer, C. Blecker, H. Gebhardt, M. Bischoff, R. Stark, K. Morgen, and D. Vaitl, "Relationship between regional hemodynamic activity and simultaneously recorded EEG-theta associated with mental arithmetic-induced workload," *Human Brain Mapping*, vol. 28, no. 8, p. 793–803, 2007.

[25] X. Hou, Y. Liu, O. Sourina, and W. Mueller-Wittig, "CogniMeter: EEG-based emotion, mental workload and stress visual monitoring," in *Proc. of International Conference on Cyberworlds (CW)*. IEEE, 2015, pp. 153–160.

[26] S. Lei and M. Roetting, "Influence of Task Combination on EEG Spectrum Modulation for Driver Workload Estimation," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 53, no. 2, p. 168–179, 2011.

[27] A. N. Patel, M. D. Howard, S. M. Roach, A. P. Jones, N. B. Bryant, C. S. Robinson, V. P. Clark, and P. K. Pilly, "Mental State Assessment and Validation Using Personalized Physiological Biometrics," *Frontiers in human neuroscience*, vol. 12, p. 221, 2018.

[28] M. A. Hafeez, S. Shakil, and S. Jangsher, "Stress Effects on Exam Performance using EEG," in *Proc. of 14th International Conference on Emerging Technologies (ICET)*, 2018, pp. 1–4.

[29] R. Sharma and K. Chopra, "EEG signal analysis and detection of stress using classification techniques," *Journal of Information and Optimization Sciences*, vol. 41, no. 1, pp. 229–238, 2020.

[30] J. A. Blanco, A. C. Vanleer, T. K. Calibo, and S. L. Firebaugh, "Single-Trial Cognitive Stress Classification Using Portable Wireless Electroencephalography," *Sensors*, vol. 19, no. 3, p. 499, 2019.

[31] M. J. Hasan and J.-M. Kim, "A Hybrid Feature Pool-Based Emotional Stress State Detection Algorithm Using EEG Signals," *Brain Sciences*, vol. 9, no. 12, p. 376, 2019.

[32] P. Nagar and D. Sethia, "Brain Mapping Based Stress Identification Using Portable EEG Based Device," in *Proc. of 11th International Conference on Communication Systems & Networks (COMSNETS)*. IEEE, 2019, pp. 601–606.

[33] M. Vourkas, S. Micheloyannis, and G. Papadourakis, "Use of ANN and Hjorth parameters in mental-task discrimination," in *Proc. of First International Conference Advances in Medical Signal and Information Processing (IEE Conf. Publ. No. 476)*, 2000, pp. 327–332.

[34] S. M. R. Islam, A. Sajol, X. Huang, and K. L. Ou, "Feature extraction and classification of EEG signal for different brain control machine," in *Proc. of 3rd International Conference on Electrical Engineering and Information Communication Technology (ICEEICT)*, 2016, pp. 1–6.

[35] Türk, M. Şeker, V. Akpolat, and M. S. Özerdem, "Classification of mental task EEG records using Hjorth parameters," in *Proc. of 25th Signal Processing and Communications Applications Conference (SIU)*, 2017, pp. 1–4.

[36] M. Yazıcı and M. Ulutaş, "Classification of EEG signals using time domain features," in *Proc. of 23nd Signal Processing and Communications Applications Conference (SIU)*, 2015, pp. 2358–2361.

[37] S. H. Oh, Y. R. Lee, and H. N. Kim, "A Novel EEG Feature Extraction Method Using Hjorth Parameter," *International Journal of Electronics and Electrical Engineering*, vol. 2, pp. 106–110, 2014.

[38] Y. Höller, J. Bergmann, A. Thomschewski, M. Kronbichler, P. Höller, J. S. Crone, E. V. Schmid, K. Butz, R. Nardone, and E. Trinka, "Comparison of EEG-Features and Classification Methods for Motor Imagery in Patients with Disorders of Consciousness," *PLoS One*, vol. 8, 2013.

[39] D. Shon, K. Im, J.-H. Park, D.-S. Lim, B. Jang, and J.-M. Kim, "Emotional Stress State Detection Using Genetic Algorithm-Based Feature Selection on EEG Signals," *International Journal of Environmental Research and Public Health*, vol. 11, 2015.

[40] M. J. Hasan and J.-M. Kim, "A Hybrid Feature Pool-Based Emotional Stress State Detection Algorithm Using EEG Signals," *Brain Sciences*, vol. 2, 2019.

[41] K. Ahammed and M. U. Ahmed, "QUANTIFICATION OF MENTAL STRESS USING COMPLEXITY ANALYSIS OF EEG SIGNALS," *Biomedical Engineering: Applications, Basis and Communications*, 2020.

[42] T. H. Priya, P. Mahalakshmi, V. Naidu, and M. Srinivas, "Stress detection from EEG using power ratio," in *Proc. of International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE)*. IEEE, 2020, pp. 1–6.

[43] F. Al-Shargie, T. B. Tang, N. Badruddin, and M. Kiguchi, "Towards multilevel mental stress assessment using SVM with ECOC: an EEG approach," *Medical & biological engineering & computing*, vol. 56, no. 1, pp. 125–136, 2018.

[44] S. Wu, X. Xu, L. Shu, and B. Hu, "Estimation of valence of emotion using two frontal eeg channels," in *2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2017, pp. 1127–1130.

[45] I. Wichakam and P. Vateekul, "An Evaluation of Feature Extraction in EEG-based Emotion Prediction with Support Vector Machines," in *2014 11th international joint conference on computer science and software engineering (JCSSE)*. IEEE, 2014, pp. 106–110.

[46] R.-N. Duan, J.-Y. Zhu, and B.-L. Lu, "Differential entropy feature for EEG-based emotion classification," in *6th International IEEE/EMBS Conference on Neural Engineering (NER)*. IEEE, 2013, pp. 81–84.

[47] D. Shon, K. Im, J.-H. Park, D.-S. Lim, B. Jang, and J.-M. Kim, "Emotional Stress State Detection Using Genetic Algorithm-Based Feature Selection on EEG Signals," *International Journal of Environmental Research and Public Health*, vol. 15, no. 11, p. 2461, 2018.

[48] S. Thejaswini, K. R. Kumar, A. Vijayendra, R. Shyam, P. D. Anchan, and E. Gowda, "An Algorithm to Detect Emotion States and Stress Levels Using EEG Signals," *International Journal of Latest Research in Engineering and Technology*, pp. 5–12, 2017.

[49] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, "Gene Selection for Cancer Classification using Support Vector Machines," *Machine Learning*, vol. 46, p. 389–422, 2002.

[50] M. Kursa and W. Rudnicki, "Feature Selection with the Boruta Package," *Journal of Statistical Software*, vol. 36, no. 11, pp. 1–13, 2010.

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

[52] H. Nahrstaedt, skjerns, D. T. Kao, S. Clarke, J. Zitting, D. Ojeda, C. Boulay, anthracene, O. Moore, M. Klumpp, D. Kosachev, and C. Miller, "holgern/pyedflib: v0.1.17," Feb. 2020.

[53] K. Dedovic, R. Renwick, N. K. Mahani, V. Engert, S. J. Lupien, and J. C. Pruessner, "The Montreal Imaging Stress Task: using functional imaging to investigate the effects of perceiving and processing psychosocial stress in the human brain," *Journal of psychiatry & neuroscience: JPN*, vol. 30, no. 5, pp. 319–325, Sep. 2005.

[54] "Montreal Imaging Stress Task - Inquisit." [Online]. Available: https://www.millisecond.com/download/library/montrealstresstest/

[55] J. I.-S. S. K. L. M.-G. Jung Ji Young, Lee Ho-Won, "Caffeine maintains arousal level and prevents change of electroencephalogram spectral powers with time at rest," *J Korean Sleep Res Soc*, vol. 11, no. 1, pp. 5–10, 2014. [Online]. Available: http://www.e-jsm.org/journal/view.php?number=171

[56] S. S. Panicker and P. Gayathri, "A survey of machine learning techniques in physiology based mental stress detection systems," *Biocybernetics and Biomedical Engineering*, vol. 39, no. 2, pp. 444 – 469, 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S020852161830367X

[57] W. K. So, S. W. Wong, J. N. Mak, and R. H. Chan, "An evaluation of mental workload with frontal EEG," *PloS one*, vol. 12, no. 4, 2017.

[58] G. Jun and K. G. Smitha, "EEG based stress level identification," in *Proc. of IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2016, pp. 003 270–003 274.

[59] D. Nath, Anubhav, M. Singh, D. Sethia, D. Kalra, and S. Indu, "A Comparative Study of Subject-Dependent and Subject-Independent Strategies for EEG-Based Emotion Recognition Using LSTM Network," in *Proc. of the 2020 the 4th International Conference on Compute and Data Analysis*, ser. ICCDA 2020. New York, NY, USA: Association for Computing Machinery, 2020, p. 142–147. [Online]. Available: https://doi.org/10.1145/3388142.3388167

[60] Anubhav, D. Nath, M. Singh, D. Sethia, D. Kalra, and S. Indu, "An Efficient Approach to EEG-Based Emotion Recognition using LSTM Network," in *Proc. of 16th IEEE International Colloquium on Signal Processing Its Applications (CSPA)*, 2020, pp. 88–92.

[61] P. Lahane, A. Vaidya, C. Umale, S. Shirude, and A. Raut, "Real time system to detect human stress using EEG signals," *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 4, no. 4, 2016.

[62] V. Vanitha and P. Krishnan, "Real time stress detection system based on EEG signals," *Biomedical Research*, pp. 271–275, 2016.

[63] Q. Wang and O. Sourina, "Real-time mental arithmetic task recognition from EEG signals," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 21, no. 2, pp. 225–232, 2013.

[64] N. Patil, M. R. P. Mirajkar, M. S. Patil, and M. P. Patil, "A Method for Detection and Reduction of Stress using EEG," *International Research Journal of Engineering and Technology (IRJET)*, vol. 04, pp. 1598–1604, 2017.

[65] M. S. Kalas and B. Momin, "Stress detection and reduction using EEG signals," in *Proc. of International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)*. IEEE, 2016, pp. 471–475.